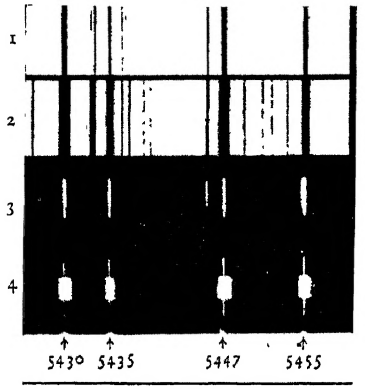
MATTER AND CHANGE

CAMBRIDGE
UNIVERSITY PRESS
LONDON: Fetter Lane
C. F. CLAY, Manager

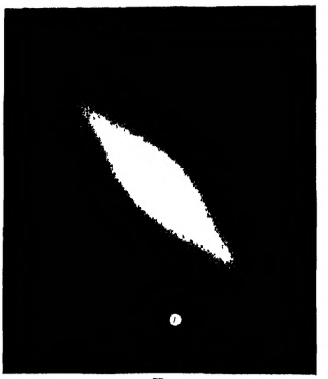
LONDON
H. K. Lewis and Co., Ltd.
Wheldon and Wesley, Ltd.
NEW YORK
The Macmillan Co.
BOMBAY, CALCUTTA and MADRAS
Macmillan and Co, Ltd.
TORONTO
The Macmillan Co of Canada, Ltd.
TOKYO
Maruzen-Kabushiki-Kaisha

All Rights reserved

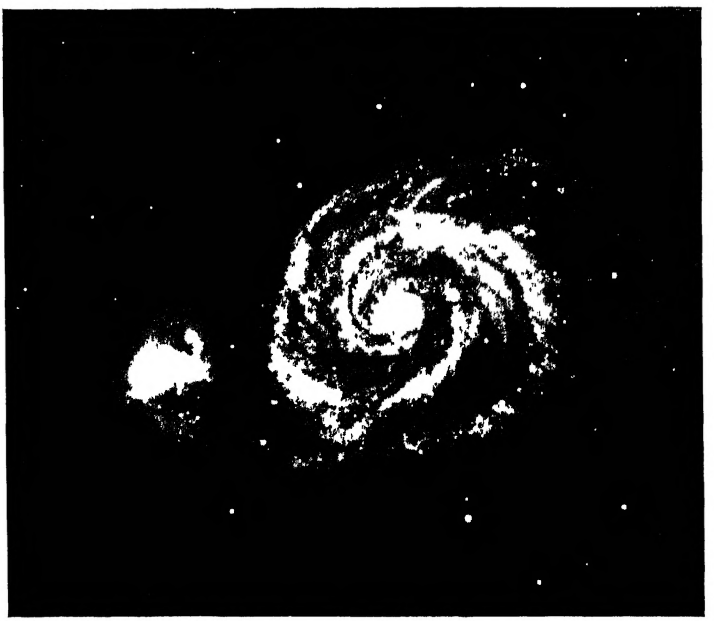
PLATE I



A



B



C

A Part of the Spectrum of (1) the Sun, (2) a Sun-spot, (3) Iron in an Electric Arc, (4) Iron in a Furnace.
B A Regular Nebula. *C* A Spiral Nebula

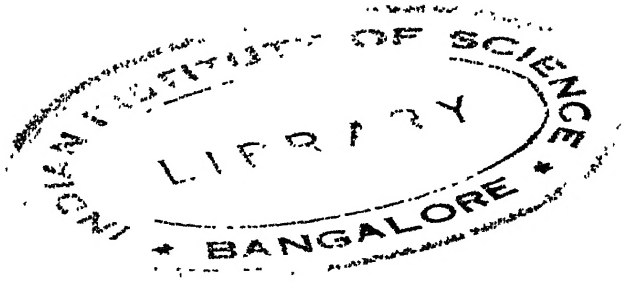
MATTER AND CHANGE

An Introduction to Physical and Chemical Science

BY

WILLIAM CECIL DAMPIER WHETHAM
M.A., F.R.S.
FELLOW AND SOMETIME SENIOR TUTOR
OF TRINITY COLLEGE, CAMBRIDGE

CAMBRIDGE
AT THE UNIVERSITY PRESS
1924



PRINTED IN GREAT BRITAIN

PREFACE

THIS book has been written at the request of the Cambridge University Press to give a short and simple statement of the methods and results of modern physical and chemical science. It is meant to be both an introduction to deeper study and an account suitable for those whose main intellectual interests lie in other fields of thought—for the upper non-scientific classes of schools and for the general reader.

I accepted the work partly because I felt that most introductory books on elementary physics and chemistry deal too much with the less interesting detail of the older part of the subjects. My aim has been to explain the essential methods and principles by illustrations which are drawn from or lead up to the more important results of recent research. I have lived through the most marvellous development of natural knowledge the world has ever seen, in touch with some of those who have made it, and I have the most profound belief that the difficulties, the failures and the triumphs of the laboratory and the observatory form some of the best material for modern literature.

Nowadays, everyone realizes the power of science as applied to industry, and all educated people have heard of the chief physical concepts by which our growing knowledge of nature is interpreted. Relativity, the electron, the complex atom, the spiral nebula whirling new stars into the depths of space, are names and ideas known unto men. But the methods by which these concepts have been won are understood by few. Even books which profess to instruct often set forth these things as though they sprang fully armed from the head of Zeus, and touch but lightly, if at all, on the fascinating tale of how men, in stress and labour, disappointment and success,

have brought them down to earth. Yet it is better to gain the spirit of scientific method, better perhaps merely to understand it, than to acquire a facile and superficial knowledge of scientific results.

Anyone who knows the simplest school-room mathematics will be able to follow the argument of this book. Where such simple tools are inadequate, I have tried to indicate the course of proofs too difficult to give at length. Always, I have first described the facts and sketched the reasoning, and so led up to the results. That is, I believe, the only true method when writing about science.

In conclusion, I wish to thank those who have lent photographs or diagrams to illustrate my text. I must also acknowledge gratefully the help both of my daughter Margaret, who wrote the section on Biochemistry, and of Miss Christine Elliott, who prepared the index.

C. DAMPIER WHETHAM

Upwater Lodge
Cambridge
Lady Day, 1924

MATTER AND CHANGE.

CORRIGENDA.

Page 21, lines 22 to 25, *read:*

"On the modern theory of the atom (chap. VIII), this atomic number gives the number of positive electric units in the nucleus, or the equal number of negative units or electrons outside it."

Page 76, line 3, *read:*

$$E = \frac{T - t}{T}.$$

Page 105, line 26, *read:*

"mechanical energy into heat..."

Page 150, line 21, *read:*

"nascent hydrogen is formed..."

Page 163, line 4, *read:*

$$\mathrm{H.CH} + \mathrm{O} = \mathrm{H.COOH}$$

Page 187, last line, *read:*

"Hence at 0° C. and..."

Page 233, lines 20 to 27, *read:*

"Other atoms consist of a complex nucleus containing positive units bound together by a less number of negative electrons, and having other electrons moving round it. *N*, Moseley's atomic number, represents the number of excess positive charges in the nucleus, or the equal number of negative electrons outside it."

Page 251, last line, *read:*

"The squares..."

CONTENTS

Chapter		*Page*
I.	MATTER	1
II.	DYNAMICS AND THE PHYSICAL PROPERTIES OF MATTER	23
III.	HEAT AND ENERGY	53
IV.	ELECTRICITY AND MAGNETISM	90
V.	CHEMICAL ACTION AND EQUILIBRIUM	124
VI.	ORGANIC AND BIO-CHEMISTRY	151
VII.	WAVES	182
VIII.	RADIOACTIVITY, ATOMIC RADIATION AND RELATIVITY	226
IX.	ASTRO-PHYSICS	249
	INDEX	271

LIST OF PLATES

PLATE I. *A.* Part of the Spectrum of (1) The sun, (2) A sun-spot, (3) Iron in an electric arc, (4) Iron in a furnace
B. A Regular Nebula
C. A Spiral Nebula . . . *Frontispiece*

PLATE II. Typical mass-spectra . . . *to face p.* 122

PLATE III. Cloud-tracks of α particles . . *to face p.* 232

Figure *A*, Plate I, is reproduced by arrangement with Messrs Edward Arnold & Co., Figures *B* and *C* by permission of Mr J. H. Jeans, F.R.S. Of the text-figures, Figures 21, 22, 28, 29, 30, 89 are reproduced by arrangement with Messrs Macmillan & Co. Ltd., Figure 56 by arrangement with Mr John Murray, and Figure 96 by permission of Mr J. Scott-Taggart.

CHAPTER I

MATTER

If we were to wake up suddenly in this wonderful world, with our intelligence fully developed but with no experience of nature, we should observe many things to which use has blunted us, and trace more easily the process of acquiring knowledge, which now has to be analyzed by laborious thought.

We should soon learn to distinguish between ourselves and not ourselves; at first, after some painful mistakes, classifying the whole of our bodies as ourselves, and afterwards, perhaps, coming to regard our bodies as in one sense outside the true sentient self, while in another they remained a real part of our being.

Then we should find that some outside things, like stones, were hard to lift, and call them heavy, while others, such as leaves or thistledown or smoke, are whirled in the breeze or float in the air and are called light. Then we might, perhaps, invent some simple apparatus and find that the property of heaviness or weight can be measured by "weighing" bodies on a pair of scales or balance.

Then we might see that heavy bodies fall quickly, while light ones take a long time to reach the ground, or even rise in the air like flame. If we merely sit and think about this fact, we may perhaps fall into the error of the great Greek naturalist and philosopher, Aristotle, and come to believe that heavy things will always fall faster than light ones and that everything has a natural "place," heavy things tending downward and light ones upward. But if, with the first of modern physicists, the Italian Galileo, we left our study and dropped a large and a small stone together from the leaning tower of Pisa or an English cliff, we should find they reach the ground together. Then, perhaps remembering that our own heavy bodies, which fall too fast for safety or comfort through the air, will float or sink very slowly through water, we might see that it is the support or resistance of the air which makes

light things fall slowly or even rise. We might perhaps arrive at the principle of Archimedes, another Greek who lived at Syracuse, who discovered that, when sunk in or floating on a fluid, a body loses a weight equal to the weight of that volume of the fluid which is displaced, so that its effective weight is by so much the less.

If we had a motor car or a bicycle, or even if we tried to walk against a gale of wind, we should learn that, besides its buoyancy, air opposes resistance to a body moving through it, a resistance which is greater the quicker the relative movement of the air and the moving body. A bicycle at first runs faster and faster down a hill. But it is opposed by a resistance increasing with its speed. And soon it travels with a steady limit of velocity: the resistance has become equal to the driving force. The weight of a rain drop depends on its volume, but this air-resistance on its diameter. Since the volume is proportional to the cube of the diameter, the resistance becomes more important the smaller the drop, and very small droplets fall so slowly that they seem to float as clouds or mist.

Solids and Fluids

Matter exists in three well-known states, solid, liquid and gas, typified by the ancients as earth, water and air.

A solid has a definite shape and a definite volume. It is often crystalline, even in cases where that is not evident: for instance, the crystals of metals may be seen on the surface of galvanized iron, that is iron coated with zinc, or by microscopic examination of polished sections cut from metal bars. If a solid be compressed, it tends to return to its original volume, and, if it be distorted, it assumes its original shape when the force is removed unless the distortion has been too great. This power of opposing a distortion is called rigidity.

Liquids and gases, grouped together as fluids, have elasticity of volume, that is to say, need force to compress them, but have no rigidity. They yield permanently to even the smallest force which tends to slide one layer slowly over the next, and have no tendency to return to their original shape when the stress is removed. This is shown by the fact that they take the shape of the vessel which contains them.

Liquids have a definite volume, but gases fill any vessel or space open to them. It needs a very large pressure to compress a liquid, a pressure of the same order of magnitude as that needed to compress a solid. Gases and vapours are much more easily compressed; indeed, as we shall see later, in ordinary circumstances, the volume of a gas like air is inversely proportional to the total pressure.

Ice can be melted into water and water evaporated as steam, and, had we complete control over all possible temperatures and pressures, it is probable that every substance likewise could be made to exist in the three states, unless chemical change occurred.

One volume of water boiled away makes about 1600 volumes of steam at atmospheric pressure, and this shows the relative order of magnitude of the volumes of liquids on the one hand and vapours or gases on the other. Most solids sink in the liquids formed by melting them, that is, they expand slightly on melting, but water forms an exception. Ice floats on water, showing an expansion on freezing, and a contraction on melting. This has an important result: it helps to prevent ponds and lakes from freezing solid even in a long frost. The ice and the cold water between 4° C. and the freezing point, being lighter than water at 4°, float on it and form a protective covering.

Inertia and Mass

If we play with a free fly-wheel and spin it round on its axis, we find that, although nothing is lifted, force must be exerted to put it in motion and to stop it when spinning. Hence we recognize another property of heavy bodies, not weight, though, as we shall find, allied to it, a property we call mass or inertia. A mass at rest remains at rest, and a mass in motion continues to move forward unless force is applied to it. We know what we mean by length and time, though recent analysis shows that these quantities depend not only on the body observed but on its relation to the particular observer. Velocity being measured in feet per second or miles per hour, its meaning is easily connected with those of length and time. From our muscular sense we know

what we mean by force, and so the experiment with a fly-wheel teaches us the idea of mass as something which needs force to set it in motion or stop it when moving.

Mass was defined by Newton as the quantity of matter in a body; and that definition is still often used. But it seems to assume that we know what we mean by matter, and, when we think about this, it appears that what we mean is something resisting us when we push or pull. Hence we see that our idea of matter really depends on our sense-perception of force. Philosophically, force seems to be a fundamental idea, because we know intuitively what we mean by a push or a pull, and cannot analyze the idea any further.

If we accept this view of things, we must consider mass to be philosophically a derived or secondary concept. If a body has a small mass, a force will give it a large acceleration, that is a high velocity in a short time. On the other hand, if the mass be large, the same force produces less acceleration. Hence we are led to measure a mass, m, by the force, f, required to produce a given acceleration, a, and may write

$$m = \frac{f}{a}.$$

The Greeks thought that force must be exerted to keep a body in motion, and this view lasted till the seventeenth century. Hence moving crystal spheres, or whirling vortices, were imagined to keep the planets travelling in their paths. But when in 1638 Galileo discovered the principle of inertia, it appeared that it was not the fact of planetary motion, but that of deflection from a straight to a circular path, that needed explanation and that a force directed to the centre would explain the deviation. Newton turned this knowledge to account and framed a law of gravity which assumed that any two particles attract each other with a force proportional to the product of their masses, and inversely proportional to the square of the distance between them. This sufficed to explain all known facts for more than two hundred years. Only the very highest degree of experimental accuracy recently attained shows that Newton's famous law may have to be replaced by a more general

explanation which connects together mass, weight and gravitation with the properties of space and time.

Weight

It is a very remarkable thing that the weight of a body is proportional to its mass, and that the most convenient method of comparing masses is by weighing them. Perhaps the reader, like all men till Newton, has been accustomed to confuse the two ideas of mass and weight. Yet there is no obvious reason to expect that the two quantities should even be proportional. As ordinarily presented they are two ideas quite independent of each other. The mass of a body means its inertia, and is measured by the force required to give it a stated acceleration. Its weight is the force which must be applied to prevent it from being drawn towards the earth. Why should the earth seem to attract a body with a force proportional to its mass? If it does so, we have the most curious fact that every particle of the body, whether on the lower surface of the body, which is nearer the earth, or inside the body, is acted on by the earth in the same way.

Yet that this proportionality between mass and weight holds good is shown by Galileo's famous experiment on the leaning tower of Pisa. Two stones of different weight let fall together reach the ground at the same moment. If the weights be written as W_1 and W_2, they represent the forces in our last equation, which therefore gives us

$$m_1 = \frac{W_1}{a_1} \text{ and } m_2 = \frac{W_2}{a_2} \text{ or } a_1 = \frac{W_1}{m_1} \text{ and } a_2 = \frac{W_2}{m_2}.$$

Galileo's experiment showed that the accelerations of the two stones are the same, or $a_1 = a_2$. Therefore

$$\frac{W_1}{m_1} = \frac{W_2}{m_2} \text{ or } \frac{W_1}{W_2} = \frac{m_1}{m_2},$$

that is, the weights of the two stones are proportional to their masses. The same result was carried to greater accuracy by Newton, who observed that two pendulums of the same length but different weights had the same time of swing.

Experimentally, then, the weight of a body is proportional

to its mass. No convincing explanation of this remarkable result was forthcoming from the time of its discovery till in 1915 Einstein applied his theory of relativity to the subject of gravitation. To that problem we shall return in the last chapter of this book.

Persistence of Matter

It is common observation that solid bodies remain unaltered over long periods of time. Those subject to use gradually wear away, but it was perceived even in early days that this could easily be explained on the supposition that minute invisible particles were removed from time to time.

Philosophers were divided into those who held it inconceivable that anything should be derived from or pass into nothing, and those who thought that everything was always changing and in a state of flux, and regarded creation or destruction as not impossible. But as soon as men began to weigh, the idea of persistence was strengthened. Bodies usually kept the same weight however they were treated, and, if divided, the sum of the weights of the parts was equal to the weight of the whole.

Chemical Action

For long ages many actions, such as burning, gave an apparent exception to this result. A torch or a candle burns away and vanishes from sight and touch. It seems as though it gave forth flame, and in so doing passed into nothingness. Since flame rises, fire was held to be something essentially light, and some of the Greek philosophers considered it as one of four elements—earth, water, air and fire.

In the eighteenth century this view was revived and elaborated by Stahl, who regarded burning as the escape of a constituent of the fuel which he called phlogiston. Since it was known that certain metals burned into an ash which was heavier than the metal, this phlogiston had to be given a negative weight, and thus chemists returned to Aristotle's conception of a body essentially light, at a time when physicists had rejected it.

The corresponding increase in weight in ordinary fuels may

be shown by the apparatus illustrated in Fig. 1. First the apparatus is weighed on a balance. The candle is then lighted, and a current of air drawn through the glass tubes. The products of combustion are absorbed by pieces of caustic soda placed in the U shaped tube. After a few minutes the candle is extinguished and the apparatus reweighed. The weight is found to have increased.

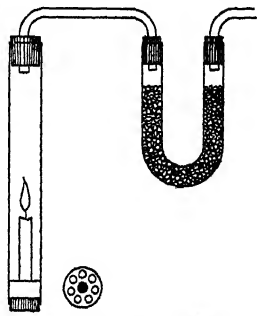


Fig. 1. Products of Combustion of a Candle.

We all know that a candle cannot burn or animals live without air, and this experiment shows that when a candle burns one or more new invisible gases must be formed which are absorbed by caustic soda. If we could have weighed the air used before and after the experiment, we should have found that it had lost as much weight as our apparatus had gained—that the total weight was unchanged.

The Beginning of Modern Chemistry

This need of air for combustion called for explanation. In 1774 Priestley discovered a new gas by strongly heating "red precipitate"—the substance obtained by gently heating mercury in air. This new gas supported combustion even more vigorously than ordinary air. Lavoisier repeated Priestley's experiment, weighing or measuring carefully all the substances concerned. He found that air was reduced by one-sixth of its volume by burning mercury in it, and that red precipitate was formed. This, heated in a small retort, gave mercury and a gas equal in volume to the loss in volume of the original air. The weight of this gas together with the weight of the mercury was equal to the weight of the red precipitate.

The simplest explanation of these results is to suppose that mercury combines with an active part of the air to form the red precipitate, which, when decomposed by further heat, gives its components again. The total weight, mass, or quantity of matter, remains constant throughout these changes.

Many of the compounds formed by this active gas have acid properties, and so Lavoisier called it oxygen (*οξύς γεννάω*).

The idea of phlogiston was now unnecessary and disappeared from science. All phenomena of combustion, and in living things the action of breathing, were explained as combinations with oxygen. Every body had a definite mass and weight, and the idea of substances essentially light with negative weights was discarded by chemists as it had been by physicists.

The gradual wearing away of solids was referred to attrition by the removal of minute, invisible particles, and the same explanation was given of the disappearance of water and other liquids. By the separation of invisible particles they pass into vapour, sometimes invisible, sometimes visible either in itself or owing to reprecipitation, as when water boils away into steam, carrying a cloud of water drops with it.

The idea of many different gases is so familiar to us that we may not realize that till modern times the only gas recognized was air, and, even in the late eighteenth century, the gases we know as oxygen and hydrogen were regarded merely as modified air.

Satisfactory experiments on gases were made possible by the invention of the pneumatic trough, which we owe to Henry Cavendish (1731–1810). A jar or tube, closed at one end and filled with water or mercury, is covered with a plate and then turned upside down with its mouth below the surface of the same liquid in a trough. The jar remains full of liquid, till a gas bubbling up through the liquid gradually fills the jar and displaces the liquid. In this way gases may be isolated from air and their properties examined. For instance, hydrogen may be produced by the action of zinc on dilute sulphuric acid in the flask shown in Fig. 2, and collected in the inverted tube over water. When the tube is full, it is easy to show that the gas has characteristic pro-

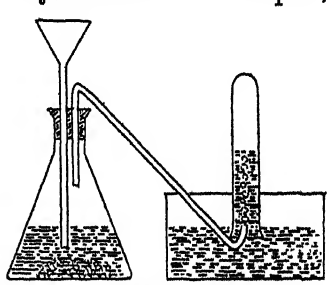


Fig. 2. Preparation of Hydrogen.

perties of its own quite different from those of air. This was first demonstrated by Cavendish in 1766. He called the gas inflammable air; the name hydrogen was given it by Lavoisier twenty years later.

Hydrogen may be poured upward from vessel to vessel, showing its lightness compared with air. When a flame is brought into contact with the gas, it burns in air, and drops of water appear.

This observation led Cavendish to the discovery that water is not an element, as all had supposed, but a compound of the two gases we now call hydrogen and oxygen. Cavendish exploded "one measure of oxygen with two measures of inflammable air" (hydrogen) in a closed vessel. No change in weight occurred, but the gases disappeared and water was formed.

The old list of four elements, earth, water, air and fire, was now seen to have no basis in fact. Visions of many chemical elements appeared—elements unknown before, and needing experiment for their discovery. In the century and a half since the days of Cavendish and Lavoisier, more than eighty have come to light.

Conservation of Matter

Priestley, Cavendish and Lavoisier founded modern chemistry, and Lavoisier showed that, in his experiments, the total weight and therefore the mass was unchanged in amount. Since then it has always been assumed that in every chemical action the total mass remains constant, and this assumption has been justified by all the countless experiments made on the strength of it. In chemical change, as in physical, matter is conserved.

We may say then that experiment shows we must imagine the concept mass, which we invent to explain inertia, to remain constant in all ordinary physical and chemical changes. If a body, such as one of the particles shot out in radioactive processes, approaches the immense velocity of light, its mass, as measured by an observer at rest, is found to increase. But, leaving on one side these very unusual cases, we find that the

mass of any isolated system of bodies is unaltered whatever changes it goes through. Whether it be left at rest or set in motion, whether it be heated or cooled, whether it be converted by chemical action between its parts into apparently quite different bodies, the total mass of the whole system is unchanged. Hence has grown up a persistent belief in the physical reality of matter. This materialist view of matter may or may not be justified philosophically, or by recent physical research. On both these grounds much may be said against it, but there is no doubt of its convenience for the purposes of elementary physics and chemistry. And in dynamics, the science of moving bodies, it is usual to take mass as a third fundamental quantity, the others being length and time, and on these three to found a system of units.

Atoms and Molecules

Even among the Greeks, two schools of philosophy held different views as to the ultimate nature of matter. One school regarded it as continuous, so that, however far divided, water was always water and earth, earth. The other school, led by Leucippus and Democritus, held that matter was composed ultimately of indivisible particles or atoms, many in size and shape but identical in substance, and that all the multitude of different properties were due to differences in size, shape, position and movement of particles of the same ultimate nature.

The Greeks possessed no experimental knowledge to enable them to decide between these two views, and the world had to wait two thousand years till John Dalton put the atomic theory on a basis which could appeal to experiment. Dalton discovered that, when two elements combine in different proportions, these proportions may be expressed in simple whole numbers. For instance, for the same weight of carbon, there is just twice as much hydrogen in one compound, marsh gas, as in another, ethylene. Again, there are several oxides of nitrogen, and in them the amounts of oxygen bear to each other simple ratios.

By the year 1803 Dalton had come to see that these relations could best be explained by the theory that the chemical

elements consist of atoms, different for different substances, but all alike for any one element, with masses which bear to each other ratios given by the combining weights. Thus Dalton found ethylene to be composed of carbon and hydrogen combined in the proportions of 5 to 1. Hydrogen being the lightest element, he took its atomic weight as unity, and this made the atomic weight of carbon 5. Writing one atom of hydrogen as H and one of carbon as C, we can represent Dalton's view by representing ethylene as CH. For marsh gas the ratio of carbon to hydrogen was found to be 5 to 2. Hence Dalton took it to contain another atom of hydrogen. This makes its formula CH_2. Changes only occur by definite steps, and that is at once explained by the theory of atoms. Again, Dalton found that one part of hydrogen combined with seven of oxygen to form water and five of nitrogen to form ammonia. Therefore he took the atomic weights of oxygen and nitrogen to be 7 and 5 respectively.

In framing his formulae, Dalton assumed that, in the compound which he thought to be the simplest, the two elements were united atom to atom. Thus he gave to water the composition HO, and to olefiant gas or ethylene the composition CH. Yet Gay Lussac and Dalton himself had found that the chemical combination of gases took place in volumes which bore simple ratios to each other. From this it follows that there must be some simple relation between the densities of gases and their chemical combining weights. The simplest hypothesis, first put forward by Avogadro in 1813 and by Ampère in 1814, is that equal volumes of all gases, whether elementary or compound, contain the same number of molecules. This led to a distinction between the molecule, the smallest mass which can exist free, and the atom, the smallest which can enter into chemical combination.

Let us take as an example the union of the gases hydrogen and chlorine, which unite explosively when exposed to bright sunlight. One part by volume of hydrogen combines with one of chlorine to form two volumes of hydrogen chloride or hydrochloric acid gas. If hydrogen and chlorine existed in the free state as single atoms, we should have

H (one volume) + Cl (one volume) = HCl (one volume).

We can explain the facts, however, at once if we assume that the molecules of hydrogen and chlorine consist each of two atoms—are diatomic. We then get

H_2 (one volume) + Cl_2 (one volume) = 2HCl (two volumes).

In the formation of water, two volumes of hydrogen unite with one of oxygen to form two of steam. Remembering that the hydrogen molecule has two atoms, we write the equation

$2H_2$ (two volumes) + O_2 (one volume) = $2H_2O$ (two volumes)

in full accordance with the facts, and see at once that water vapour is not HO, as Dalton supposed, but H_2O, two atoms of hydrogen combined with one of oxygen. Most ordinary gases consist of diatomic molecules, but a few, such as argon and mercury vapour, are monatomic.

Water is the only stable compound of hydrogen and oxygen, and its composition shows that one oxygen atom naturally combines with two atoms of hydrogen. This is expressed by saying that the "chemical valency" of oxygen is two. Hydrogen and chlorine are monovalent atoms, oxygen and sulphur are divalent, aluminium and boron, which most naturally combine with three monovalent atoms, are trivalent. Sometimes an element has two valencies: iron, for instance, can be either divalent or trivalent, phosphorus trivalent or pentavalent.

Thus, by experiments on the volumes of gases which react, the atomic weights of elements that enter into gaseous combination may be estimated. If neither the element nor its compounds are gaseous, the method cannot be applied, and a choice between the possible values has to be made on the best other evidence available. When a vapour condenses to a liquid or solid, as steam to liquid water, the proportion between the atoms in the molecule must remain unchanged, but combination between molecules may occur. Thus, while water vapour is H_2O, liquid water may be n (H_2O), or $H_{2n}O_n$, where n is any whole number. There is some evidence to show that the molecule of liquid water is 8 (H_2O), and that some solid crystals may be regarded as very complex single molecules.

Avogadro's hypothesis, when its importance was appreciated, served to complete Dalton's atomic theory, and to

start the measurements of atomic weights once for all on the right lines. The hypothesis is confirmed by the fact that it may be deduced from what is called the kinetic theory of gases, a theory which naturally developed when it was understood that heat was a form of energy, the energy of motion of the molecules of bodies. To this theory we shall return in the chapter on Heat, where we shall also give some calculations of the number of molecules in a given volume. The success of the kinetic view of heat also lent additional support to the general atomic theory, which may be said to have been finally established by Dalton's writings and the clear description of the kinetic view of matter, first given in 1844 by James Prescott Joule.

Metals

Of the different classes of elements which make up all the types of matter known to us, one of the most important to mankind is that of the metals. Metals have been known from early times. Tools of bronze supplanted those of stone and flint, and gave way in turn to those made of iron.

A few metals are found in the metallic state in nature. Gold, for instance, occurs as veins in quartz, and can be dissolved out by mercury or cyanides when the quartz is crushed under stamps. It is also found in river sands and gravels and can be separated by washing. Platinum, too, is found in similar places. Gold and platinum are examples of metals very difficult to corrode or dissolve in acids, and this gives them their special value, and explains their occurrence in the free state in nature.

Other metals more easily acted on chemically, such as copper and silver, occur both in the metallic and in the combined forms. Their existence as free metals explains their early use.

Their chief ores are oxides and sulphides. The oxides may be reduced by roasting them with charcoal or coke in a furnace. The carbon of the fuel, or the carbon monoxide formed by burning it, unites with the oxygen of the ore to form oxides of carbon, and the metal is liberated. Sulphides

are usually first roasted in air to convert them into oxides, and then treated as described.

The extraction of copper from cuprite or ruby ore, Cu_2O, may be taken as an example of the simplest operation. The air-blast burns the carbon to carbon monoxide:

$$2C + O_2 = 2CO,$$

and this gas reduces the copper oxide to copper:

$$Cu_2O + CO = 2Cu + CO_2.$$

Iron, being easily oxidized and dissolved in acids, is seldom found in the metallic state in nature. The simplest ore is red haematite, ferric oxide, Fe_2O_3. It is reduced in a blast furnace, a long cylindrical tower into the bottom of which a blast of hot air is blown. The top is fed with a mixture of the ore, coke and limestone. The reactions are complicated, but the chief one is the reduction of most of the ferric oxide by carbon monoxide

$$Fe_2O_3 + 3CO = 2Fe + 3CO_2.$$

As the charge descends to the hotter parts of the furnace, any oxide left will react with carbon

$$Fe_2O_3 + 3C = 3CO + 2Fe.$$

The iron absorbs or dissolves carbon and small quantities of other substances such as silicon, phosphorus and manganese, and collects as a molten mass at the bottom of the furnace with a quantity of slag floating on it. The iron is run off at intervals.

Commercial iron always contains carbon, and usually other substances. Its physical properties depend largely on these constituents: cast iron may contain about 3 per cent. of carbon, steel about one-half per cent. and wrought iron less than one-tenth of 1 per cent.

Metals more easily oxidized than iron are difficult to extract and are of most importance as compounds. Thus calcium as carbonate is the substance of chalk and limestone, aluminium as silicate forms clay, and sodium as chloride gives us our common salt.

To mankind the metals which are most important in the state of elements are iron and copper for the practical arts,

and gold for its use as money. All heavy engineering products and tools of all kinds depend on iron, and electrical industry needs enormous quantities of copper. Gold is the best standard of value if a metallic currency is used. It has a value for other purposes; it is non-corrodable, and therefore nearly permanent, so that the amount in circulation is large compared with that which is produced in any one year. Nevertheless, variation in the demand for and supply of gold is responsible for some of the greatest social and economic changes in history. The discovery and exploitation of the New World in the sixteenth century, or of the Australian and South African gold mines in the nineteenth century, by increasing the world's supply of gold, decreased its value compared with other commodities, that is, caused steadily rising prices and consequent eras of agricultural and trade prosperity. On the other hand, the increased demand for gold from 1871–1895, when many countries were adopting a gold standard for the first time, sent up its value. Prices measured in gold consequently fell during these years; more especially in agriculture, owing to the simultaneous development in ocean transport. Many farmers were ruined, though some other industries were better able to adjust themselves and suffered less severe periods of depression. On the other hand, those who possessed or earned fixed incomes gained by the fall in prices, while wages either did not fall or at all events not in the same proportion as wholesale commodities. Owing to the development of modern systems of credit, the actual supply of gold is now able to carry a much larger volume of trade at a given price-level, while, since 1914, few countries have kept a gold standard. Nevertheless, the rate of output and movements of gold still have a great economic importance.

Acid-forming Elements

Sulphur or brimstone is found free in nature, and has been known from the earliest historical times. It has played a great part in the history of chemistry, and for a long while was regarded as "the principle of fire."

It easily burns in air, combining with the oxygen to form the gas sulphur dioxide, SO_2. This gas may be oxidized

further by contact with oxides of nitrogen to sulphur trioxide or SO_3, which with water, H_2O, gives sulphuric acid, H_2SO_4—the so-called oil of vitriol, a dense, viscous, intensely corrosive liquid.

Sulphuric acid reacts with many metals such as zinc or iron to form salts called sulphates. In the preparation of hydrogen, described on page 8, if the liquid left in the bottle be evaporated, transparent colourless crystals are deposited from the solution. They contain zinc, sulphur and oxygen in the proportions represented by $ZnSO_4$, more loosely combined with seven molecules of water to form $ZnSO_4.7H_2O$. The water may be driven off by gently heating the crystals.

Other sulphates of natural or commercial importance are ammonium sulphate, used as an artificial manure owing to the nitrogen in its ammonia; sodium sulphate or Glauber's salt; calcium sulphate, which in the hydrated form of $CaSO_4.2H_2O$ is known as gypsum, and of $(CaSO_4)_2\,H_2O$ as Plaster of Paris; and double sulphates of aluminium with potassium, sodium or ammonium known as the alums.

If common salt be heated with sulphuric acid, a colourless heavy gas having the composition HCl is evolved. This gas fumes strongly in air and colours litmus red. It is the same substance which is obtained by combining hydrogen and chlorine in sunlight, as described on page 11. It has the composition HCl and is called hydrogen chloride. It is very soluble in water; 1 cubic centimetre of water dissolves 525 c.c. of the gas at 0° C. and 440 c.c. at 20° C. This solution is called hydrochloric acid.

If heated with oxidizing agents, such as manganese dioxide, hydrochloric acid evolves a yellowish-green gas, somewhat soluble in water, very irritating and poisonous if breathed—the first "poison gas" used in war. When discovered by Scheele in 1774 many people thought it was a compound, but efforts to decompose it failed, and it was finally recognized as an element and called chlorine (Cl) by Davy. The gas of hydrochloric acid, as we have seen, is a compound of chlorine and hydrogen, HCl, and the extraction of chlorine consists in the oxidation of hydrogen to water

$$4HCl + MnO_2 = MnCl_2 + 2H_2O + Cl_2.$$

Chlorine was the first of the gases thought to be permanent which was liquefied. It forms a compound chlorine hydrate with water, and if this be heated in one end of a sealed bent tube (see Fig. 3), chlorine is evolved and collects as liquid in the other end if it be immersed in ice.

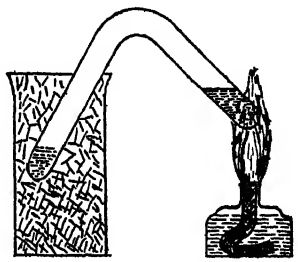


Fig. 3.

Hydrochloric acid dissolves many metals to form chlorides, as sulphuric acid forms sulphates. Important chlorides are common salt, NaCl, calcium chloride, $CaCl_2$, and calcium oxychloride or bleaching powder, Ca(OCl)Cl.

Nitrogen, which forms four-fifths by volume of the air, in ordinary circumstances is a very inert gas, but, by passing a current of air through an electric arc light, the nitrogen may be combined directly with oxygen. From the oxides so formed, nitric acid, HNO_3, may be obtained. It is more easily produced by the action of sulphuric acid on nitrates.

Nitric acid has a very intense oxidizing action. It will dissolve all metals except a few like gold and platinum, and even these yield to a mixture of nitric and hydrochloric acids known as *aqua regia*. Nearly all nitrates are soluble in water. They are of very great industrial importance.

Nitrates are essential to the fertility of the soil, where they are produced by growing leguminous crops, which have the power of fixing nitrogen from the air, or are added in natural or artificial manures. The latter have hitherto been chiefly obtained from nitrate deposits in Chile, but these beds are not inexhaustible, and an alternative source is very desirable.

One successful method, largely used in Scandinavia, is a development of that indicated above—a blast of air is blown through an electric arc. Nitrogen and oxygen from the air combine with each other, and from the oxides thus formed nitrates are obtained.

Nitrates are also used for most explosives, and during the war the artificial production of nitrates for this purpose as well as for fertilizers was stimulated greatly in Germany by the blockade.

In particular, in a process chiefly due to Haber, nitrogen and hydrogen are made to combine directly to form ammonia. The process is carried out at very high pressures of some 2000 atmospheres and at a temperature of 500° to 700° C. A catalyst (see chap. v) is employed, though its nature has not been divulged. The ammonia may be oxidized to nitrates by the action of oxygen in presence of platinum.

Countless other mineral acids are known, but the most important are these three—sulphuric, hydrochloric and nitric.

Salts

A metallic oxide or base unites directly with an acid to form a salt. Acids are sour to the taste and turn blue litmus red. Some metallic hydrates and carbonates, formed from the oxides, are alkaline—they are soapy to the touch and turn red litmus blue. The salts formed by the union of an acid with an alkali or base are neutral in all these ways. As an example caustic soda and hydrochloric acid give water and sodium chloride

$$NaOH + HCl = NaCl + H_2O.$$

If this reaction takes place in solution in a large quantity of water, we shall later see reason to suppose that the base, the acid and the salt exist chiefly as dissociated, oppositely electrified particles, which, since they travel under the influence of an electric current, are called ions. The neutralization of an acid by an alkali is merely the formation of water from its ions, and may be written as

$$\overset{+}{Na} + \overset{-}{OH} + \overset{+}{H} + \overset{-}{Cl} = \overset{+}{Na} + \overset{-}{Cl} + H_2O$$

This interpretation is borne out by the fact that the heat evolved by these neutralizations is independent of the nature of the reagents, is, for instance, the same in the above named reaction as in the neutralization of potash by nitric acid

$$\overset{+}{K} + \overset{-}{OH} + \overset{+}{H} + \overset{-}{NO_3} = \overset{+}{K} + \overset{-}{NO_3} + H_2O.$$

The solutions which result deposit solid crystals of salt on evaporation.

Such salts as these are readily soluble in water, but all

degrees of solubility are known, some salts being practically insoluble. The formation of insoluble salts serves as the best test for many metals and acids in analysis. Thus silver chloride is deposited as a white insoluble precipitate when the solution of a soluble silver salt such as the nitrate is mixed with the solution of a soluble chloride such as common salt. Hence chlorides serve as a test for silver, and silver nitrate can be used as a test for chlorides. In quantitative analysis the mass of insoluble silver chloride is weighed, and thus the amount of silver or chlorine calculated.

The determination of acids and alkalies may be more easily made by volumetric analysis or titration. The solution of (say) acid of unknown concentration is coloured with litmus or some other indicator. A standard solution of an alkali of known strength is then run or dropped slowly in from a graduated glass tube, and a note made of the volume used when the colour just changes. This gives the amount of alkali needed to neutralize the acid, and thus the concentration of the acid is measured.

The action of the indicator depends on the liberation of a coloured ion. Hence the exact meaning of neutralization depends on the indicator, slight differences occurring for instance, when litmus is replaced by methyl orange or phenolphthalein.

Carbon and its Compounds

Of all chemical elements the most remarkable is carbon. It exists as charcoal and also as crystal diamond, probably formed under immense pressure and at very high temperature.

It forms two oxides: carbon monoxide, CO, when it burns in a limited supply of oxygen, and carbon dioxide, CO_2, when combustion is complete.

Carbon dioxide dissolves in water, the quantity increasing in proportion to the pressure, a fact which is made use of in the manufacture of aerated waters. The liquid acts as a weak acid, which forms a series of well marked salts, the carbonates. They include such well known substances as washing soda, $Na_2CO_3\,10H_2O$, that is, sodium carbonate crystallized with ten molecules of water, and marble, chalk and limestone, which are all calcium carbonate, $CaCO_3$,

formed in different physical conditions. Calcium carbonate is nearly insoluble in water, but dissolves in excess of carbon dioxide to form an acid carbonate, $H_2Ca(CO_3)_2$. When this so-called "hard" water is boiled, carbon dioxide is driven off, and the insoluble carbonate is precipitated in our kettles and boilers. Another acid carbonate, $HNaCO_3$, is known in commerce as bicarbonate of soda.

The most important property of carbon is its power of combining with itself and other atoms to form all the numberless complex bodies known in organic chemistry and in living matter. These will be considered in a future chapter. The extensive modern sciences of organic and bio-chemistry are but the beginnings of a study of the carbon compounds.

The Periodic Table

It was observed by Newlands and by Mendeleeff that if the elements be arranged in a table in the order of ascending atomic weights, elements with similar physical and chemical properties come at recurring intervals. Thus one such group contains Lithium, Sodium, Potassium, etc., and another contains Fluorine, Chlorine, Bromine and Iodine. The properties are dependent on the atomic weights, and the weights therefore are clearly quantities of very fundamental importance. Indeed in this way we are brought back to the Greek idea of a common basis for matter, the idea that all different kinds of matter are made by the combinations of particles of one primordial substance. The fact that many of the atomic weights approach whole numbers when hydrogen is taken as unity, or still more nearly when oxygen is taken as 16, suggested to Prout that all the elements were made up of atoms of hydrogen. For long this hypothesis broke down on such cases as that of chlorine, the atomic weight of which is 35·46, but recently it has been shown by F. W. Aston that many of these discordant elements are made up of mixtures of atoms of identical chemical properties but of different atomic weights, chlorine, for example, being a mixture of atoms of weights 35 and 37. Again, lead is the final product of some radioactive changes (see chap. VIII), and Soddy has found that lead prepared from radium has an atomic

weight of 206 instead of 207·2. Thus ordinary lead is probably a mixture. These elements, with identical chemical properties, but different atomic weights, are called isotopes.

In the past, whenever a new physical or chemical method of research has been applied to the problem, new elements have been discovered. Sir Humphry Davy isolated potassium and sodium from potash and soda, formerly thought to be elements, when in 1806 he passed through them Volta's new electric currents. The first satisfactory periodic table arranged by Mendeléeff showed by gaps the existence of undiscovered elements. The invention of spectrum analysis, described in chapter VII, revealed many others. Work on radioactivity disclosed radium, polonium, actinium, etc.

The most important advance in this subject since the discovery of the periodic table itself was made by H. J. G. Moseley in 1913 and 1914, by a research on X-ray spectra (chap. VII). Moseley showed that the square roots of the frequency of the characteristic X-rays increase by equal steps as we ascend the list of elements, so that each element can be given a number which at once expresses its place in the periodic table and a physical quantity of great significance. On the modern theory of the atom (chap. VIII), this atomic number gives the number of positively electric hydrogen units in the nucleus, or the total number of negative electric charges or electrons in the atom. The atomic numbers in the table of elements which follows have then this real physical meaning, and the periodic table acquires a new importance.

We now (1924) know 87 elements, ranging in weight from hydrogen 1, to uranium 238·2, and in atomic number from 1 to 92. We must not forget that some of these elements are mixtures of isotopes with the same number but different weights. There also exist certain radioactive products with the same atomic numbers as elements in the table, such as radium-C with the 83 of bismuth, radium-D with the 81 of lead, and ionium with the 90 of thorium.

The table which follows shows five gaps, and we are therefore justified in concluding that there are still five undiscovered elements between hydrogen and uranium.

Table of Elements

Atomic Number	Element	Symbol	Atomic Weight
1.	Hydrogen	H	1·008
2.	Helium	He	4·00
3.	Lithium	Li	6·94
4.	Beryllium	Be	9.1
5.	Boron	B	11·0
6.	Carbon	C	12·005
7.	Nitrogen	N	14·01
8.	Oxygen	O	16·00
9.	Fluorine	F	19·0
10.	Neon	Ne	20·2
11.	Sodium	Na	23·00
12.	Magnesium	Mg	24·32
13.	Aluminium	Al	27·1
14.	Silicon	Si	28·3
15.	Phosphorus	P	31·04
16.	Sulphur	S	32·06
17.	Chlorine	Cl	35·46
18.	Argon	A	39·88
19.	Potassium	K	39·10
20.	Calcium	Ca	40·07
21.	Scandium	Sc	44·1
22.	Titanium	Ti	48·1
23.	Vanadium	V	51·0
24.	Chromium	Cr	52·0
25.	Manganese	Mn	54·93
26.	Iron	Fe	55·84
27.	Cobalt	Co	58·97
28.	Nickel	Ni	58·68
29.	Copper	Cu	63·57
30.	Zinc	Zn	65·37
31.	Gallium	Ga	69·9
32.	Germanium	Ge	72·5
33.	Arsenic	As	74·96
34.	Selenium	Se	79·2
35.	Bromine	Br	79·92
36.	Krypton	Kr	82·92
37.	Rubidium	Rb	85·45
38.	Strontium	Sr	87·63
39.	Yttrium	Y	88·7
40.	Zirconium	Zr	90·6
41.	Niobium	Nb	93·5
42.	Molybdenum	Mo	96·0
43.	—	—	—
44.	Ruthenium	Ru	101·7
45.	Rhodium	Rh	102·9
46.	Palladium	Pd	106·7
47.	Silver	Ag	107·88
48.	Cadmium	Cd	112·40
49.	Indium	In	114·8
50.	Tin	Sn	118·7
51.	Antimony	Sb	120·2
52.	Tellurium	Te	127·5
53.	Iodine	I	126·92
54.	Xenon	Xe	130·2
55.	Caesium	Cs	132·81
56.	Barium	Ba	137·37
57.	Lanthanum	La	139·0
58.	Cerium	Ce	140·25
59.	Praseodymium	Pr	140·9
60.	Neodymium	Nd	144·3
61.	—	—	—
62.	Samarium	Sm	150·4
63.	Europium	Eu	152·0
64.	Gadolinium	Gd	157·3
65.	Terbium	Tb	159·2
66.	Dysprosium	Ds	162·5
67.	Holmium	Ho	163·5
68.	Erbium	Er	167·7
69.	Thulium	Tm	168·5
70.	Ytterbium	Yb	173·5
71.	Lutecium	Lu	175·0
72.	Hafnium	Hf	—
73.	Tantalum	Ta	181·5
74.	Tungsten	W	184·0
75.	—	—	—
76.	Osmium	Os	190·9
77.	Iridium	Ir	193·1
78.	Platinum	Pt	195·2
79.	Gold	Au	197·2
80.	Mercury	Hg	200·6
81.	Thallium	Tl	204·0
82.	Lead	Pb	207·2
83.	Bismuth	Bi	208·0
84.	Polonium	Po	210·0
85.	—	—	—
86.	Niton (Radium emanation)		222·0
87.	—	—	—
88.	Radium	Ra	226·0
89.	Actinium	Ac	—
90.	Thorium	Th	232·0
91.	Proto-actinium	Pa	—
92.	Uranium	U	238·2

CHAPTER II

DYNAMICS AND THE PHYSICAL PROPERTIES OF MATTER

Fundamental Quantities

As we have seen, there are certain fundamental conceptions, such as length and time and force, which at present we need not seek to explain in any simpler terms. They arise in our minds from the direct perception of the world by our senses.

Space

The idea of length seems to come as the simplest conception of space from our senses of touch and sight. We look and feel along the length of a stick, and, compared with other sticks, call it long or short.

When we ask how long, we get the idea of measurement, and, to carry out this idea, we need some unit in which to measure. Probably the yard was originally a man's easy span or arm's length, but at an early stage of civilization it became desirable to have a fixed or standard yard, though we are still not civilized enough to have constant standards for all the quantities we need to measure.

A standard yard was made by direction of Parliament in 1760, but was lost when the Houses of Parliament were burned in 1834. A new standard was made in 1844, and since then the legal yard has been the length, at a definite temperature of 62° Fahrenheit, between two scratches on a metal bar kept at the Standards Office of the Board of Trade in London. In France the unit is the metre, which, in conformity with the systematizing instinct of the first French Republic, was originally meant to be the one ten-millionth part of a quadrant, that is, the quarter of a circumference of the earth. The measurement of the earth on which the original metre was made has, of course, been superseded by later and more accurate determinations, but the metre has not been altered accordingly; it remains the length laid off on a platinum bar

at Paris, at the end of the eighteenth century, and is now as arbitrary a unit as is the British yard.

Feet, inches, centimetres and kilometres are secondary units, derived from the fundamental units of the yard and the metre.

Metric Units of Length

A kilometre is 1000 metres.
A metre is the standard.
A decimetre is 0·1 m., that is $\frac{1}{10}$ of a metre.
A centimetre is 0·01 m., ,, $\frac{1}{100}$,,
A millimetre is 0·001 m., ,, $\frac{1}{1000}$,,
A metre is 39·37 inches or 1 inch = 2·540 cms.

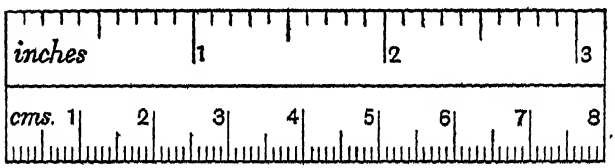


Fig. 4.

Next we grasp the idea of surface, the simplest limited surface being a square, with four sides and four angles equal. Then volume—the cube, a solid of six square faces, with all edges and angles equal.

Thus surface and volume are ideas which can be derived from that of length—a square yard or metre and a cubic yard or metre being the appropriate units. The cubic metre is too large for a laboratory unit of volume, so a cubic decimetre, the thousandth part of a cubic metre, is taken as a unit and called a litre. One thousandth part of a litre is a cubic centimetre, written as 1 c.c.

Thus from rough sense perceptions the mind conceives space in three dimensions—length, breadth and thickness. Whether this corresponds with the facts of actual space to an unlimited order of accuracy remains a matter of inquiry and observation. However this may be, our idea of surface is the square of a length, and of volume the cube of a length. This is sometimes expressed by saying that, if the "physical dimensions" of length are written as L, those of surface are L^2 and those of volume are L^3.

Time

The next fundamental and unexplainable idea is that of time—a sequence in our states of consciousness. The mind most easily conceives time as independent of space, and constructs an imaginary scheme of things in which space is that described by Euclid, and the mechanics are those of Newton, who imagined time as independent of space or man, flowing on with absolute uniformity. Once again, whether this sense of space, time and their derivatives is in exact accordance with nature remains a subject for experiment. As a matter of historical fact it has been found to agree with nature till the very highest order of accuracy was reached in quite recent experiment, when indications have appeared that the real space-time is not quite what we believed. If we come to think (which often leads to unexpected results) there is no obvious reason to expect that what seems simplest to our minds after a superficial examination of nature must correspond accurately with reality. The structure of the Universe is not necessarily adapted to the schedule of an elementary examination.

The unit of time is fixed by astronomical observations. The solar day is defined as the time between two successive passages of the sun across the meridian, the vertical plane through the observer and the geographical north and south poles of the earth. The day varies at different times of the year, and consequently the mean solar day—the average length of the solar day throughout the year—is chosen as the standard unit. The mean solar day is divided into 24 hours, each hour into 60 minutes, and each minute into 60 seconds. Thus one second is the one 86,400th part of the mean solar day, and one second is taken as the practical laboratory unit of time.

Velocity and Acceleration

We can now consider moving bodies. Speed or velocity brings in the ideas of both length and time, and is measured, for example, by so many feet per second or miles per hour, the unit velocity in the first case being that of one foot per second. Therefore the physical dimensions of velocity are

those of length divided by time. We may put this in the form of an equation

$$[v] = \frac{L}{T},$$

$[v]$ being written to denote the physical dimensions of v, a velocity, and L and T to denote the fundamental units of length and time.

When the speed of a body is changing, the rate of increase of speed is called the acceleration, and is measured by the velocity added in unit time. The dimensional equation is therefore

$$[a] = \frac{[v]}{T} = \frac{L}{T^2},$$

which is conveniently written as $[a] = LT^{-2}$.

If we are using feet and seconds as our units of length and time, the unit of acceleration will be an acceleration of one foot per second per second, that is, an extra velocity of one foot per second added in one second to the moving body.

Mass and Inertia

As explained in the first chapter, the idea of force may be taken as fundamental, and that of mass derived from it. A large force is needed to impart a given acceleration to a large mass, and thus the mass of a body may be defined as the force needed to give it unit acceleration. As thus defined, mass has been shown (see page 5) to be proportional to weight, and by means of the balance Lavoisier established the principle of the conservation of matter. This property of persistence makes mass a convenient quantity to take as a third fundamental quantity on which to base our system of dynamical and physical units. For this purpose, the theoretical reasons in favour of treating force as fundamental and mass as derived, are overridden by practical convenience. Mass is taken as the third basal unit, and force treated as a secondary unit derived from it.

It was Galileo who first investigated the effects of inertia, which depends on mass as distinguished from weight. Till his day, it was thought that a continual exertion of force was needed to keep a body in motion. Galileo allowed a ball to roll down an inclined plane, and found that it would just reach the top

of another plane of equal vertical height whatever was the length of the second plane. By imagining the length of the second plane extended, it can be seen that the ball will roll farther and farther, and, if the plane be made horizontal, it will move indefinitely with steady speed, except as friction gradually and slowly brings it to rest.

Newton formulated Galileo's discovery in his first law of motion—every body continues in a state of rest or of uniform motion in a straight line except when acted on by an impressed force.

Of late years it has become clear that this statement involves much assumption. We have no direct experience of a body moving for ever in a straight line with no force acting on it. A stone sliding on ice, or Galileo's ball rolling along a plane, are upheld by the constraining force exerted by the ice or the plane. If this supporting force were removed, they would fall towards the earth in a curved path, as do projectiles. On the face of it, the bodies when falling freely are less constrained than when forced to move over the ice or the plane. The ideas that the earth attracts the bodies, and that this gravitational force is exactly balanced by the upward thrust of the ice or the plane, are assumptions based on Newton's theory of gravitation. If we clear these assumptions from our minds, we must regard the actual paths taken by a falling stone or a projectile, or indeed by a planet revolving in an orbit about the sun, as the natural paths of those bodies, and must study them as such, instead of regarding as a natural path a hypothetical straight line, and seeking for hypothetical forces to explain any deviation from it. We shall find in chapter VIII that this simple and direct way of taking the actual paths as the natural paths of moving bodies leads to a revolutionary change in outlook, and gives a complete explanation of the facts of gravitation, with no need to invoke Newton's mysterious and ever inexplicable force of attraction.

But, for the moment, let us suppose a straight line to be the natural path of a body, as indeed it is when no other body is near. This is the simplest case, and we can deal with other cases later.

A moving body can only be stopped by the exertion of force. We are familiar with the devastating effect of opposing a heavy mass in motion. There is clearly some important quantity which depends on both mass and velocity. The product of these two, $m \times v$, is called "momentum," and the change in momentum when the body is brought to rest is called "impact."

Units of Mass

The British unit of mass is the pound, a certain lump of platinum kept at the Standards Office in London. It bears no definite relation to our units of volume, though it happens that a gallon of water weighs almost exactly ten pounds. One pound (written as lb.) is equal to 7000 grains or 453·593 grams.

The French unit of mass, the gramme or gram (written as gm.), was originally meant to be the mass of one cubic centimetre of water at 4° Centigrade, and this is very nearly accurate, though the gram is now legally the one thousandth part of a lump of platinum named a kilogram, and kept in Paris.

Metric Units of Mass

One kilogram, the standard, is 1000 gms.
One gram is the nominal standard, but is really the one thousandth part of a kilogram, the actual standard.
One decigram is 0·1 gm., that is $\frac{1}{10}$ of a gram.
One centigram is 0·01 gm., „ $\frac{1}{100}$ „
One milligram is 0·001 gm., „ $\frac{1}{1000}$ „

For almost all scientific purposes, measurements are based on the centimetre, the gram and the second as units. This is called the C.G.S. system.

Force

As we have seen, since mass is a quantity which shows constancy, it is well to take it as a practical fundamental unit. On this practical system force becomes a secondary or derived unit, in spite of the philosophical reasons for treating it as fundamental. A force is measured by the acceleration it will impart to unit mass, that is, by the rate of change of momentum it causes. This statement is equivalent to Newton's second law of motion. Our three fundamental dynamical

quantities being taken as length L, time T, and mass M, the physical dimensions of force are given by the equation

$$[f] = [ma] = \frac{ML}{T^2} = MLT^{-2}.$$

If our units are the pound, the foot and the second, the unit of force will be that which will give to a pound an acceleration of one foot per second per second. This unit force is called the poundal. The corresponding French unit will give to a gram an acceleration of one centimetre per second per second, and is called the dyne.

Work and Power

Another familiar dynamical idea is that of work. When a horse pulls a cart, or an engine pumps water or drives machinery, we say that work is being done. If we lean against a firm wall, we are doing no external work, though the mere exertion of force seems to involve physiological changes inside our bodies. To get mechanical work done, the applied force must move the body it is applied to, and we can define the amount of work done, W, as the measure of the force multiplied by the measure of the length through which it moves its point of application. The consequent dimensional equation is $[w] = [f] \times L = ML^2T^{-2}$.

The French unit of work, the erg, is the amount of work done when the force of one dyne is exerted through the distance of one centimetre. The correct British unit of work should be the foot-poundal, but a more common unit is the foot-pound, the work done when the weight of one pound is lifted one foot.

Power is the rate of doing work, that is, the amount of work done per unit time—for instance, the number of ergs of work done per second. The most usual practical unit of power is the horse-power, which was meant to be about the work which one horse could do, and is defined as 33,000 foot-pounds per minute.

Energy

The amount of work which a body or a system can do is called its energy. That form of words seems at first sight satisfactory, but we cannot estimate the total amount of

work which a body could do if we might draw on the internal energy of its atoms and molecules. Indeed, the ordinary work available will depend on the position of the body with regard to other bodies. A two pound weight raised on to a shelf five feet from the floor can do ten foot-pounds of work in falling. But, if the floor be at the top of a building, more work can be obtained by letting the weight fall to the ground outside, and more again by dropping it down the shaft of a mine. Hence we see that we must confine ourselves to changes in total energy, and be content to say that the amount of work a body will do in passing from one state to another measures the energy it gives out during that change.

Mechanical energy can exist in two forms: kinetic in virtue of mass in motion, which will do work in being brought to rest; and potential in virtue of position—for instance, in clock-weights wound to a height which will drive the clock while they fall.

A body reaches a state of equilibrium when its potential energy is a minimum—a stone when it is at rest on the ground, a pendulum when its string is vertical and its bob at the lowest point. This tendency of potential energy to decrease to a minimum gives a convenient way of calculating the conditions of equilibrium in many cases.

We cannot do work without expending an equivalent amount of energy, and we shall see in future chapters that the total amount of energy in an isolated system is constant and unchangeable in amount—that energy, like matter, is conserved. Indeed we are just beginning to see some reason to think that matter itself may be a form of energy.

Falling Bodies

Galileo also discovered the laws of falling bodies. As before, he reduced the speed to measurable values by rolling balls down inclined planes instead of letting them fall freely through the air, and, after one false start, he guessed the right relation and confirmed his hypothesis experimentally. The velocity of a falling body increases steadily with the time. If the body is falling freely, it is found to increase in speed by

32·2 feet per second, or 981 centimetres per second, in each second of time; that is, the acceleration (g) due to gravity is uniform and equal to 32·2 feet or 981 centimetres per second per second, and, as Galileo observed, this is the same for all bodies heavy enough for the effect of air resistance to be neglected.

The results of these experiments can be expressed in shorthand thus. The velocity v of a falling body at the end of t seconds from rest is

$$v = gt \qquad \text{.....(1),}$$

so that the velocity at the end of 1 second is 32 feet per second, at the end of 2 seconds 64, at the end of 3 seconds 96, and at the end of 4 seconds 128, and so on till the speed becomes so great that the resistance of the air reduces appreciably the rate of increase.

Since the velocity increases uniformly, its average value is $\frac{1}{2}gt$. The space (s) traversed with a uniform velocity would be vt, and so, if the average velocity is $\frac{1}{2}gt$,

$$s = \tfrac{1}{2}gt \times t = \tfrac{1}{2}gt^2 \qquad \text{.....(2).}$$

Hence in 1 second a body falls through 16 feet

2	seconds	,,	,,	64	,,
3	,,	,,	,,	144	,,
4	,,	,,	,,	256	,,

and so on, these numbers being in practice decreased by a successively increasing amount on account of the resistance of the air. When the body moves so fast that the resistance of the air is equal to the weight of the body, no resultant force acts on it, and it then falls with a uniform velocity.

Since $s = \frac{1}{2}gt^2$ and $t = \dfrac{v}{g}$,

we have $s = \frac{1}{2}g \times \dfrac{v^2}{g^2} = \dfrac{1}{2}\dfrac{v^2}{g}$,

and $$v^2 = 2gs \qquad \text{.....(3).}$$

Equations (1), (2) and (3) give the complete laws of falling bodies.

Projectiles

If a ball be thrown horizontally from the top of a cliff it moves forward with uniform velocity (neglecting air resistance) and at the same time falls vertically, as described above, with a velocity increasing as the square of the distance. Its resultant path is a curve called a parabola, and its position

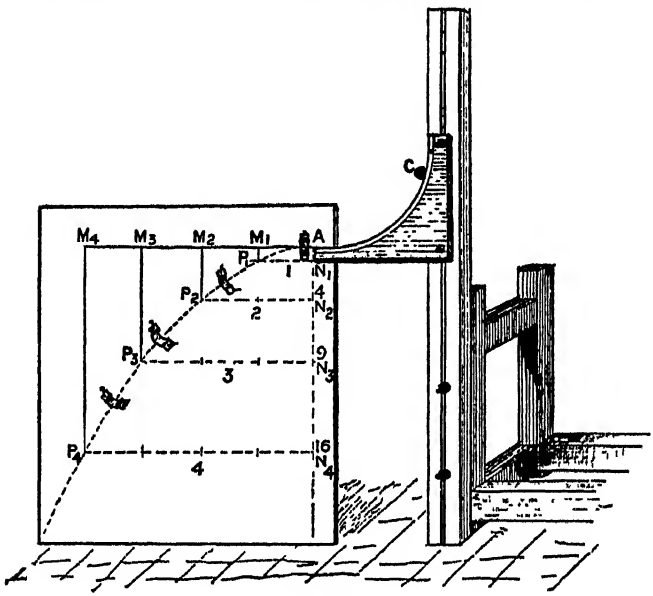


Fig. 5.

at any time may easily be calculated by treating its two motions as independent. The motion may be investigated experimentally by allowing the ball c (Fig. 5) to roll down the grooved board so that it is projected horizontally at A, and fixing paper loops by trial and error till it passes through them each time. It will be found that the vertical distances M_1P_1, M_2P_2, M_3P_3, etc. are proportional to the squares of the horizontal distances AM_1, AM_2, AM_3, etc. This is the characteristic property of the parabola.

If, instead of horizontally, it be thrown vertically upwards

its velocity is reduced by a uniform negative acceleration and finally reversed. It returns to earth with its original velocity. If it be projected at an angle, the original velocity v, represented by AC in Fig. 6, can be resolved into a vertical velocity v_1, and a horizontal velocity v_2. Neglecting for a moment the effect of gravity, if AC represent v, the space the body would describe in one second if it continued to move with its initial velocity, it will reach C at the end of the second, and will have moved vertically over a space equal to AB and horizontally over one equal to AD. Thus AB and AD, the sides of the rectangle AC, represent v_1 and v_2, the resolved velocities.

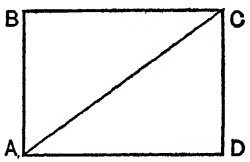


Fig. 6.

Now, with a real projectile, while the horizontal velocity is unaffected by gravity, the vertical velocity is first diminished and then reversed with uniform acceleration. We can calculate the flight of the projectile, therefore, by treating these velocities independently, v_2 being constant, and v_1 subject to a uniform acceleration in the opposite direction to that of the original velocity. Again, if we neglect the resistance of the air, we get a parabola.

The Parallelogram Law

The resolution of velocity into two components serves to introduce and illustrate a general law. If a body move in one second from O to B (Fig. 7) over, say, 3 feet, and at the same time the line OB move over 4 feet into the position AC, the actual path of the body will be OC, the diagonal of the parallelogram OC which, with the angles we have chanced to draw, is about 6 feet. Hence the resultant velocity is the diagonal of the parallelogram of which the sides are the resolved velocities.

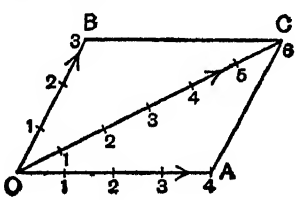


Fig. 7.

If the body move with acceleration, we may take OB and OA to represent the velocity added in these directions in one

second, that is, the accelerations. Therefore accelerations may be resolved or compounded by the same process as velocities, and, since accelerations are proportional to forces, the same law applies, and we get a parallelogram of forces also.

Hence we see that any two forces of which the lines of action meet at a point, may be replaced by a resultant force also passing through the point. But we are left with the case of two parallel forces, of which the lines do not meet. What is the resultant of the two forces P and Q (Fig. 8), or, what comes to the same thing, what force opposite to the resultant must be applied to keep P and Q in equilibrium?

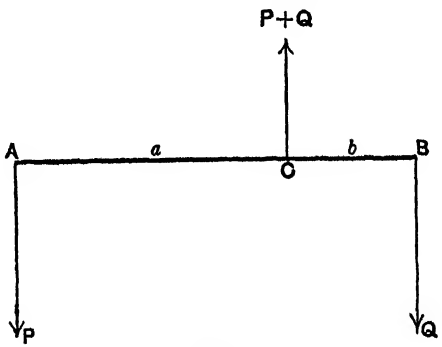


Fig. 8.

If P and Q be applied at A and B at the opposite ends of a straight bar, suspended or balanced over a knife edge or fulcrum at C, we see that the problem is that of the lever. This can best be solved experimentally. It is found by observation that, if the weight of the bar itself is small enough to be neglected, the bar will balance when the length AC or a is to the length BC or b as the force Q is to the force P, or $Pa = Qb$.

Work and Energy

If b is made short and a long, we get the ordinary lever used to raise a heavy weight at B by a moderate force applied at A. If the weight is raised through a height h_2, the work done is the force Q multiplied by the distance moved h_2, or Qh_2. Similarly the work done at the other end is Ph_1. By similar

triangles h_1 is to a as h_2 is to b, and since $Pa = Qb$, we get $Ph_1 = Qh_2$, that is, the work done by the force applied at A is equal to the work done on the weight raised at B.

The same result holds good for all machines—other kinds of lever, pulleys, wheels and axles, etc. The amount of work done on the load is equal to the energy expended by the force applied to move the machine if no loss occurs from friction, and less than the energy expended if such losses are taken into account.

Since a moving body continues to move till brought to rest by a force, it will do work, and therefore possesses energy called kinetic energy by virtue of its velocity, the energy being measured by the amount of work done by the body in coming to rest.

Galileo's laws of falling bodies apply to all bodies subject to a uniform acceleration α. If a force f be opposed to the motion of a mass m moving with a speed v, the work done is fs, where s is the space the mass moves against the force before being brought to rest. The force is $m\alpha$, and so,

$$w = fs = mas,$$

α is the velocity added or taken away in unit time, or $\frac{v}{t}$, and s is the average velocity, that is, half the initial velocity, multiplied by the time, or $\frac{1}{2}vt$. Hence

$$w = m \times \frac{v}{t} \times \tfrac{1}{2}vt = \tfrac{1}{2}mv^2,$$

and $\frac{1}{2}mv^2$ measures the work done, that is, the energy of the motion of the body relative to other bodies. It is called kinetic energy.

If the force which brings the body to rest is reversible and not frictional, if, for instance, it is exerted by a spring or by gravity on a body shot vertically into the air, fs measures the potential energy when the body is brought to rest. It may be regarded as stored in the system by virtue of the position of the body relative to others.

Here the potential energy fs gained is equal to the kinetic energy $\frac{1}{2}mv^2$ expended. If the body move back again under the force, the work done or the potential energy lost is equal to

the kinetic energy acquired. Again, we find that the total energy is unchanged; what is gained in kinetic energy is lost in potential. Mechanical energy, at all events, is conserved.

Rotation

The experimental principle of the lever is not only of fundamental importance in statics or the science of the equilibrium of forces, but, conjoined with the equally fundamental experiments of Galileo, opens a new field in dynamics.

The important quantity in the lever is clearly the product of the force P multiplied by the length of the arm a at which the force acts round the fixed centre c (Fig. 8, p. 34). This product is known as the moment of the force. If the moments of two forces about a fixed point are equal and opposite, the system is in equilibrium. If not, it begins to revolve about the fixed point. If two equal forces tend to turn the system in the same direction, the sum of their moments is called a couple.

Let us imagine the fly-wheel of an engine or the balance wheel of a watch to consist essentially of a heavy rim of mass m at a distance t from the centre of revolution or oscillation, the mass of the spokes being negligible. If the wheel be rotating with a velocity of circumference v, its kinetic energy is $\frac{1}{2}mv^2$.

The angular velocity ω, is the angle in circular measure (that is, the arc divided by the radius), described per second and is therefore equal to v/r. Thus $v = r\omega$, and the kinetic energy is $\frac{1}{2}mr^2\omega^2$.

Now mr^2 is a constant quantity, dependent only on the dimensions and mass of the wheel. It is called the moment of inertia and may be written as K.

The work done by a couple c of two equal forces acting at the circumference is $2f \times l$, where l is the distance through which a point on the circumference is moved. The couple c is $2fr$ or $2f = c/r$ and the work done is

$$c\frac{l}{r} = c\theta,$$

where θ is the angle through which the wheel moves.

For a linear movement we have (see page 35)

$$fl = \tfrac{1}{2}mv^2,$$

and we see that for circular motion we get a corresponding equation

$$c\theta = \tfrac{1}{2}mr^2\omega^2 = \tfrac{1}{2}K\omega^2.$$

We can now dispense with the simplification of a wheel with a heavy rim, and imagine a body of any shape, rotating about an axis. The body can be supposed made up of a very large number of particles each of mass m. The distances r of the particles from the axis will vary, but, if the body be rigid, the angular velocity of them all about the axis will be the same. Hence, if we add together all the quantities mr^2 for the different particles, and write the sum as $\Sigma\,(mr^2)$, we can put the last equation as

$$c\theta = \tfrac{1}{2}\Sigma\,(mr^2)\,\omega^2 = \tfrac{1}{2}K\omega^2,$$

and see at once that the moment of inertia of a body of any form about a fixed axis is $\Sigma\,(mr^2)$ and that, in circular motion, this quantity corresponds to m, the mass in linear motion.

Harmonic Oscillations

A light coiled steel spring, suspending a heavy weight of mass m, gives the simplest case of oscillatory motion (see Fig. 9). It is easy to show by experiment that the force which the spring exerts is proportional to its extension, so that, if we pull down the weight and let it go, during its up and down vibrations it is acted on by a force tending to bring it to its position of equilibrium and at each point proportional to the distance of that point from the point of equilibrium.

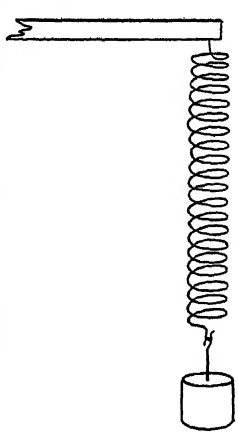


Fig. 9.

To investigate the result we must begin by considering a case of circular motion (see Fig. 10). Let a point P describe a circular orbit about a centre C. The velocity of P is uniform and is always directed along the circumference at right angles to the radius CP. Hence, if we draw CV at right angles to CP, and cut off a length CV

proportional to the velocity, the position of CV shows the velocity of P from moment to moment.

The acceleration of P is a change in direction, and not in magnitude, of the velocity, and clearly takes place along the radius PC. Now, just as velocity is rate of change of position, so acceleration is rate of change of velocity. Therefore acceleration bears to velocity the same relation as velocity to position. CA, then, drawn at right angles to CV and in length a third proportional to CP and CV, represents the acceleration of P.

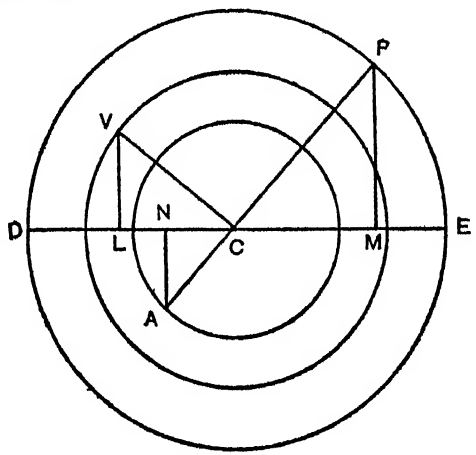


Fig. 10.

Let us imagine a mass M to travel backwards and forwards along the diameter DE so that it always lies at the foot of the perpendicular drawn from P to DE. It is then said to travel with harmonic motion. Those with a little knowledge of trigonometry will see that the displacement of M from C undergoes periodic changes which vary as the cosine of the steadily increasing angle PCM, but that knowledge is not necessary to our argument.

As CM represents the displacement of M as P revolves steadily in its orbit, so CL and CN, where the perpendiculars from V and A cut DE, represent the resolved velocity and acceleration of P or the actual velocity and acceleration of M.

Since the triangles PCM and ACN are similar,

$$\frac{CN}{CM} = \frac{CA}{CP}$$

and CA and CP, the radii of the circles, are constant. Therefore $CN = CM \times$ constant, or the acceleration α of M is proportional to its displacement d from C. But force is proportional to acceleration, and so the force acting on M is proportional to its displacement, and the body M, moving with harmonic motion, is subjected to a force like that exerted by the spring in Fig. 9.

Now CA is a third proportional to CP and CV, so

$$\frac{CP}{CV} = \frac{CV}{CA} \text{ or } \frac{r}{v} = \frac{v}{CA} \text{ and } CA = \frac{v^2}{r}.$$

This gives the acceleration of the body P moving in a circular orbit, and, since force is mass acceleration, the force which we may imagine acting towards the centre to keep the body (whirling stone, moon or earth) in its orbit is

$$F = \frac{mv^2}{r}.$$

This is the so-called centrifugal or centripetal force.

Now, returning to the diagram, since, as above,

$$\frac{CA}{r} = \frac{CN}{CM}$$

we have

$$\frac{CN}{CM} = \frac{\alpha}{d} = \frac{CA}{r} = \frac{v^2}{r^2}.$$

The time T taken by M to oscillate from D to E and back again, that is, to perform one complete oscillation, is the same as the time needed for P to describe its circular orbit, a time which is measured by the length of the circumference $2\pi r$ divided by the velocity v.

Hence

$$T = \frac{2\pi r}{v} = 2\pi\sqrt{\frac{d}{\alpha}} = 2\pi\sqrt{\frac{1}{\alpha_1}}$$

where α_1 is the acceleration of M for unit displacement. But force = mass $\times$ acceleration, hence we may write the last equation as

$$T = 2\pi\sqrt{\frac{m}{F_1}}$$

where F_1 is the force when the displacement is unity. Thus the time of oscillation is independent of the amplitude, that is, does not depend on the length of the swing—the vibrations are said to be isochronous.

This result gives the time of oscillation for a linear oscillation, such as is given by the weight and spring in Fig. 9. When a body is oscillating about an axis, the corresponding relation is

$$T = 2\pi \sqrt{\frac{K}{C_1}}$$

where K is the moment of inertia, and C_1 the couple when the angular displacement is unity.

This last expression shows that a watch will keep time if the balance wheel is controlled by a spring which exerts on the wheel a couple proportional to the angle through which the wheel is turned. The oscillations are then isochronous—each takes the same time.

The pendulum of a clock may be treated either as an instance of circular motion about the axis of suspension or of linear motion of the bob. When the pendulum is pulled aside to a position OP (Fig. 11) the vertical force on the bob is its weight mg, and the component of this tending to restore the bob to its position of equilibrium is

$$mg \times \frac{AP}{OA}.$$

As long as the angle of displacement is small this is equal to

$$mg \frac{AP}{OP}.$$

While this is so, the force is proportional to AP or CP, and therefore to the displacement, so that the vibrations are isochronous. Their period is given by

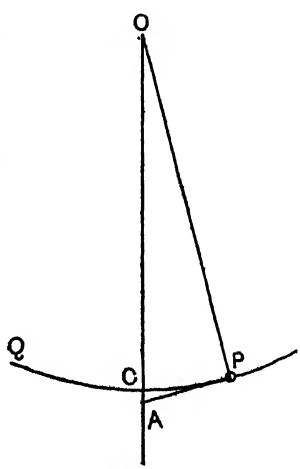


Fig. 11.

$$T = 2\pi \sqrt{\frac{m}{F_1}}.$$

F_1 is the force along AP divided by the displacement AP, or

$$F_1 = mg\frac{AP}{OP \times AP} = \frac{mg}{OP} = \frac{mg}{l}$$

where l is the length of the string OP.

Thus we get

$$T = 2\pi\sqrt{\frac{m}{\frac{mg}{l}}} = 2\pi\sqrt{\frac{l}{g}},$$

and see that as long as the swing is small, the time of swing does not depend on the amplitude of the swing or on the mass of the bob, but only on the length of the string and the acceleration g of gravity. Hence a pendulum is used to measure the value of g, which varies slightly from place to place on the earth's surface.

Once more we owe the discovery of the pendulum to Galileo. Watching the swing of a lamp in the cathedral of Pisa, he saw that its period was the same for small as for larger swings, and in the light of this discovery he invented the pendulum clock.

Elasticity

A fluid, whether liquid or gaseous, has no rigidity, and gives way permanently under a shearing stress—that is a uniform force per unit area on one layer tending to slide it over the next, as each card in a pack may slide a little over the one below it and the pack lean over at one end. On the other hand, a fluid resists compression, that is to say, it has elasticity of volume. If p be the pressure applied, V the original volume and v the decrease in volume produced, the elasticity E is defined by the relation

$$E = \frac{p}{v/V}.$$

A solid has volume elasticity of the same order of magnitude as a liquid. It also possesses rigidity: it opposes a shearing stress tending to change its shape, and, if the strain has not been too great, recovers its original shape when the stress is removed. If the shearing stress is F per unit area and if that stress changes the angle of a cube of the solid by θ, the rigidity n is given by

$$n = \frac{F}{\theta} = \frac{F}{a/r},$$

θ being expressed in circular measure, in which an angle is measured by the arc a over the radius r.

Hydrostatics

The fact that fluids have no rigidity gives them qualities which are studied under the subject called hydrostatics. The fundamental discovery was made by Archimedes, who, noticing the rise in the level of the water when he got into his bath, saw that a solid sunk in a fluid displaces its own volume of fluid, and loses in weight by an amount equal to the weight of fluid displaced.

Density and Specific Gravity

This principle of Archimedes leads at once to the idea of density or specific gravity. The density of a substance is defined as the mass per unit volume, in the C.G.S. system of units, the number of grams per cubic centimetre. The specific gravity is the ratio of the weight of a substance to the weight of an equal volume of water.

The specific gravity of a solid is found by weighing the solid first in air, and then in water as shown in Fig. 12. In

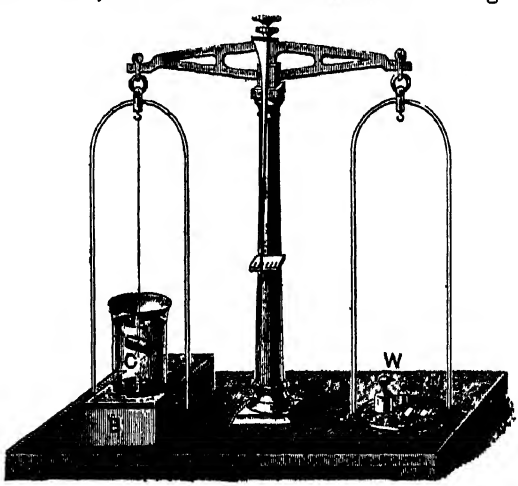


Fig. 12.

the first case the solid is buoyed up by the air which surrounds it, but so are the weights, and the necessary correction is small and can usually be neglected. When weighed in water, the observed weight W_2 is $W_1 - W$, where W_1 is the weight of the solid in air and W is the weight of an equal volume of water. Hence the specific gravity S_1 is given by

$$S_1 = \frac{W_1}{W} = \frac{W_1}{W_1 - W_2}.$$

The specific gravity of a liquid may be found by weighing this same solid in that liquid. If its weight then be W_3, the weight of an equal volume of the liquid is $W_1 - W_3$ and the ratio of this to the weight of an equal volume of water, that is, the specific gravity S_2 of the liquid is

$$S_2 = \frac{W_1 - W_3}{W_1 - W_2}.$$

The density of water is best determined by weighing in air and then in water a solid of regular form such as a carefully turned cylinder. The dimensions of the cylinder are measured accurately with calipers, and its volume V calculated. The density of the water is then equal to

$$\frac{W_1 - W_2}{V}.$$

From this result the densities of other solids and liquids may be found as described above.

The density of air or of another gas is less easily determined. A round bottomed glass flask (Fig. 13) is exhausted by an air pump, hung from the scale pan of a balance and carefully weighed. First dry air and then the gas to be examined is allowed to enter to a known pressure and each time the flask is again counterpoised on the balance. The increase in weight gives the mass of air and of the gas which has entered. The volume of the globe may be found by filling it with water and weighing it full and empty on a rougher, stronger balance.

Fig. 13.

A vapour is a gas which can be condensed

to a liquid by pressure. Its density may be determined by the apparatus shown in Fig. 14.

A glass bulb is drawn off to a narrow neck (at O) and weighed, W_1. A small quantity of the volatile liquid, the vapour of which is to be examined, is then put in the bulb, and the bulb sunk in a hot bath above the temperature at which the liquid boils. When vapour ceases to escape from the bulb, the tip of its tube is quickly sealed in a blowpipe flame. The bulb is then removed from the bath, allowed to cool and weighed a second time, W_2. The top is broken off,

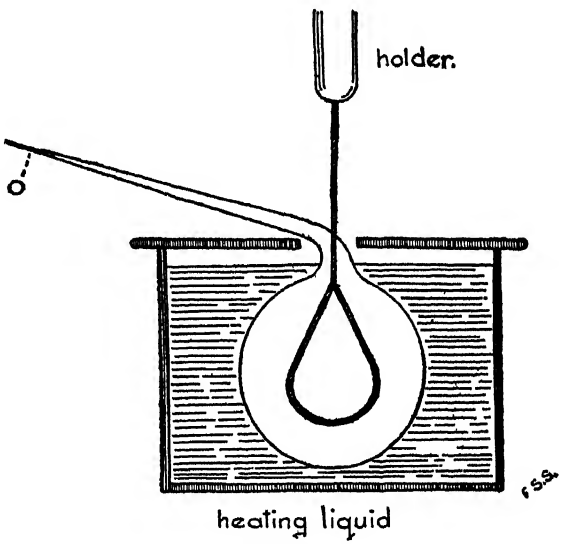


Fig. 14.

and the bulb filled with water and once more weighed, W_3. The weight of water which fills it is $W_3 - W_1$, and this gives its volume V in cubic centimetres. The weight W of this volume of air can be calculated from its known density, and the weight of the bulb exhausted would be $W_1 - W$. The weight of the vapour which filled the bulb is $W_2 - (W_1 - W)$ and its density $= \dfrac{W_2 - (W_1 - W)}{V}$.

By Avogadro's hypothesis, which was considered in chapter 1, equal volumes of gases at the same temperature and pressure contain the same number of molecules. Hence a determination of the density of a gas or vapour gives its molecular weight. Hydrogen, the lightest of gases, is taken as a standard. At the standard temperature and pressure, 0° C. and 760 mm. of mercury, 1 gram of hydrogen fills 11·16 litres or 11,160 cubic centimetres. The molecule of hydrogen contains two atoms (see page 12) so its molecular weight is 2 and its gram-molecule, i.e. 2 grams, fills 22,320 c.c. From these figures and a density determination, the relative densities referred to hydrogen taken as 2 may be tabulated as molecular weights:

Hydrogen ...	2·0	Nitrogen	28·0
Oxygen ...	32·0	Water vapour ...	18·0
Chlorine ...	71·0	Hydrogen chloride ...	36·5

The atomic weights of the diatomic elementary gases hydrogen, oxygen, chlorine and nitrogen are half these numbers. Water vapour, having a density of 18, must consist of two hydrogen atoms 2 + one oxygen atom 16 = 18, and hydrogen chloride of one hydrogen 1 + one chlorine 35·5 = 36·5. This agrees with the results described on pages 11 and 12 and shows how density determinations throw light on chemical composition.

The Pressure of Fluids

The old saying that water finds its own level expresses one case of the general principle that a fluid is driven from a place where the pressure is higher to one where it is lower; it moves in the direction of the force upon it. Pressure is defined as the force per unit area; so the pressure on the bottom of a vessel of water is the sum of (1) the weight of water divided by the area of the bottom, which gives the weight standing on unit area, and (2) the pressure of the atmosphere, which weighs on the top of the water and is transmitted through it to the bottom. In two vessels connected by a tube the pressure below a higher column will be greater than that below a lower column, and the liquid will move until the levels are the same. If the connecting tube be very narrow and the friction great, the levels slowly adjust themselves, but, if the connecting tube

be wide, the liquid will oscillate backwards and forwards several times before coming to rest, though in the end the levels will be identical, that is, the pressure everywhere at the same level in a fluid must be the same.

The Barometer

The pressure of the atmosphere was first measured by Torricelli, a pupil of Galileo, who invented the barometer in 1643. A glass tube closed at one end is filled with mercury, and then inverted over mercury in a trough. If the tube is more than 30 inches in length, the mercury in it will fall, leaving a space above it, till it stands about 30 inches above the surface of the mercury in the trough. At this level the pressure on the surface of the mercury outside the tube is that of the atmosphere, and at the same level inside the tube is that due to the column of mercury above it. The pressures anywhere at the same level must be equal, and so the height of the column of mercury in the tube above the mercury in the trough is a measure of the pressure of the atmosphere. If the pressure of the atmosphere increases, the height of the mercury in the tube increases also, perhaps to 31 or 31·5 ins., while in stormy weather, when pressure is low, it may fall to 29 inches or lower. The average height is about 30 inches, and 30 inches or 76 cms. of mercury is taken as the standard pressure of the atmosphere.

A cubic centimetre of mercury weighs 13·596 grams at the freezing point of water, and so a column 76 cms. high presses on a square centimetre at its foot with a force equal to the weight of 1033 grams. Since weight is the force due to gravity, and is the mass multiplied by the acceleration, it is mg, so that, to convert the weight into the c.g.s. units of dynes per square centimetre, we must multiply by the acceleration of gravity, 981. Hence in c.g.s. units, the pressure of the standard atmosphere is $1{\cdot}014 \times 10^6$, or a little more than one million dynes per square centimetre. This corresponds to about 15 pounds weight per square inch in British units.

Since liquids are almost incompressible, the pressure in a liquid increases regularly with the depth. But a layer of air at the surface of the earth is compressed by the atmosphere

above it, and a cubic centimetre weighs more than it does at the top of a mountain.

Boyle's Law

The compressibility of air was investigated by Robert Boyle by means of an apparatus the same in principle as that shown in Fig. 15. The closed end of the tube at A is filled with dry air. The total pressure p_1 on it is the pressure of the atmosphere plus the pressure due to the difference in level of the mercury in the two tubes B and C. At this pressure the air occupies a volume equal (say) to v_1 c.c. The tube C is now raised; the volume contracts to v_2, when the total pressure (atmosphere and mercury column) is p_2. Boyle found that the volume was inversely proportional to the pressure or

$$p_1 v_1 = p_2 v_2,$$

and this relation is called Boyle's Law. It is found to hold good very accurately for air and similar gases, but less accurately for gases such as carbon dioxide or chlorine, more easily condensed to liquids.

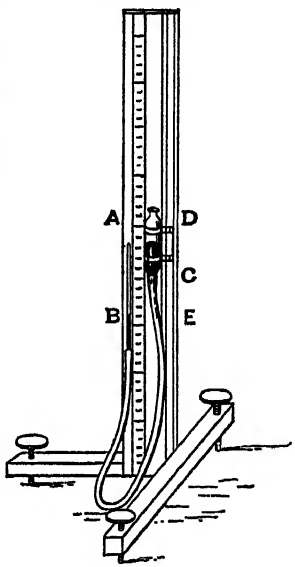


Fig. 15.

Thus, as we rise in the atmosphere, not only is there a less height of air above us, but that air weighs less per unit volume. Hence the pressure falls in a geometrical progression. If we go up 1000 feet, the barometer falls from 30 to 29 inches of mercury, that is, by one-thirtieth part of its height at sea level. Another 1000 feet rise again causes the mercury to fall, but this time by one-thirtieth of 29, that is, by $\frac{29}{30}$ inch, so that at 2000 feet it stands at $29 - \frac{29}{30}$ or $28\frac{1}{30}$ inches. With this law to help us, we can calculate heights by observing the fall in a barometer as we ascend.

Boyle's Law enables us to find a numerical value for the volume elasticity of a gas. If the pressure be increased slightly

from P to $P+p$ and the volume consequently decreased from V to $V-v$, we have

$$PV = (P+p)(V-v)$$
$$= PV - Pv + pV - pv.$$

Now p and v are small quantities, so their product, pv, may be neglected. Thus

$$PV = PV - Pv + pV,$$

or

$$Pv = pV,$$

and

$$E = \frac{p}{v/V} = P.$$

That is, the volume elasticity of a gas at constant temperature is numerically equal to its pressure. At one atmosphere pressure and at 0° C. the isothermal elasticity of air is therefore about 10^6 in C.G.S. units.

Surface Energy or Surface Tension

If water be dropped from a fine nozzle, the drops look quite round, and the phenomena of the rainbow when considered mathematically show that rain drops are very perfect spheres. The sphere is the figure which has the smallest surface for a given volume, and this indicates that, when falling freely, that is, when removed from disturbing external forces, the forces inherent in the drop tend to bring its surface to a minimum. A system is in equilibrium when its potential energy is at a minimum, and thus we see that part of the potential energy of one body surrounded by air or another substance increases with the area of contact, and comes to a minimum when the surface is a minimum. It is found that all phenomena agree with the supposition that the total surface energy is proportional to the area of surface, so that we may define a useful quantity, S, as the surface energy per unit area, that is,

$$S = \frac{\text{energy}}{\text{area}}.$$

Now an energy (ML^2T^{-2}) per unit area (L^2) has dimensions MT^{-2}, the same as those of a force per unit length

$$\left(\frac{MLT^{-2}}{L} = MT^{-2}\right).$$

Hence surface energy may also be regarded as surface tension, and the rain drop as enclosed in a stretched elastic skin. Indeed we can see by common sense that such a skin will produce the same effect and round the drop into a sphere.

If a small quantity of oil be placed on the surface of water, it spreads out into a film. On the other hand, if water be placed on an oily or greasy surface, it gathers into drops. Here we have three media, water, oil and air, and we have to deal with three surface energies: that between water and air, S_{wa}; that between water and oil, S_{wo}; and that between oil and air, S_{oa}.

When oil spreads over water, it increases the area of contact between water and oil and also that between oil and air, while it diminishes the surface between water and air. Since the potential energy must tend to a minimum, this shows that

$$S_{wa} \text{ is greater than } S_{wo} + S_{oa}.$$

Hence, when water is placed on oil, we know that it cannot spread so as to cover the oil-air contact, for, from the above equation, the potential energy is decreased by making the water-air surface as small as possible, that is, by gathering the water together into drops, even though some of the water is thus raised above the surface of the oil against the force of gravity. Indeed from the thickness of such a drop the surface tension may be calculated.

When we are dealing with large masses, the surface energy is usually very small compared with the other energy involved, but if matter is finely divided, as in drops of spray, or liquid in the crannies of a porous mass, or emulsions of particles scattered throughout a liquid solvent, the surface energy may be the determining factor in physical or chemical equilibrium. The physics and chemistry of particles called colloids is now of great importance. Colloids play a great part in the physiology of plants and animals, and the special properties of the soil in agriculture are due to the presence of colloid particles, as well as of living organisms such as bacteria and amoebae. The soil is not the simple mixture of mineral salts with decaying vegetable matter we used to think it. It has a structure and life of its own, and on that

structure and life its fertility depends. Again, the presence of moisture in the soil depends on another effect of surface tension—the rise of liquids in narrow chinks and in capillary tubes.

If a thin glass tube with a bore of a millimetre or less be held vertically, and its lower end dipped in water, the water will rise in the tube. Here again we have three surfaces—one between glass and air, one between glass and water, and one between water and air. In equilibrium the potential energy is a minimum, which means that, for a small change in height, the energy will be unaltered (see chap. v). The increase in energy due to raising in the tube a thin layer of thickness l of the water through a height h against gravity is mgh, and, since the mass m is the volume $\pi r^2 l$ multiplied by the density d, $mgh = \pi r^2 ldgh$, and this must be equal to the loss in energy due to surface tension, $2\pi rl\,(S_{ga} - S_{gw})$.

Hence
$$h = \frac{2\,(S_{ga} - S_{gw})}{rdg}.$$

Water wets glass, and so the angle of contact at the upper surface of the water is zero, and, since there is equilibrium between the three surface forces, we must have

$$S_{gw} + S_{wa} = S_{ga} \text{ or } S_{ga} - S_{gw} = S_{wa}.$$

Thus, if we write S for S_{wa}, the surface tension between water and air, we get
$$h = \frac{2S}{rdg},$$

where r is the radius of the tube, d the density of water and g the acceleration of gravity. This result gives the easiest way of measuring surface tension. The radius of a capillary glass tube is estimated, and the height to which the liquid rises observed.

It will be seen that the narrower the passages through a spongy mass the higher the water will rise. A finely divided soil such as clay will retain moisture better than a coarser soil like sand or gravel.

Since the total potential energy of a surface is proportional to its area, as we have seen, the area of a free surface between two media tends to decrease. An interesting result follows if we apply this principle to a collection of water drops in a

space saturated with aqueous vapour. The water-air surface is greater the more the liquid is broken up—that is, the smaller are the drops. The drops will tend to become fewer and larger, the larger ones growing by condensation of vapour as the small ones diminish by evaporation. It follows that the vapour pressure of the small drops must be greater than that of the large ones, the pressure increasing with the curvature of the surface. The exact relation may easily be found by considering the equilibrium of its vapour with water risen in a capillary tube where the upper surface is concave, and in the vessel at the foot of the tube where the vessel is lower and and the pressure greater.

Viscosity

As we have seen, fluids oppose a resistance to bodies moving through them, a resistance which becomes greater the greater is the velocity. It depends also on the nature of the fluid, the resistance offered by water being greater than that of air, and that of treacle greater than that of water. This property of fluids is called viscosity.

Suppose that two flat plates are placed parallel to each other at a small distance apart in a liquid. Let the lower one be fixed and the upper one moved forward slowly parallel to itself. The force tending to drag the lower plate after the upper one, is found to be proportional to the area of the plates, to their relative velocity, and inversely to the distance between them:

$$F = \frac{Av}{d} \times \text{constant}.$$

This constant is called the coefficient of viscosity and is usually written as μ.

$$\mu = \frac{Fd}{Av},$$

and its physical dimensions are

$$[\mu] = \frac{MLT^{-2} \times L}{L^2 \times L/T} = ML^{-1}T^{-1}.$$

We may use this result to illustrate the practical importance of the theory of dimensions in the treatment of problems where the general nature of the solution is clear. A rain drop falling through air suffers a retarding force which might

depend on the viscosity μ of the air, on the radius r of the drop, and the velocity v of fall. Let us suppose it to be proportional to powers of these quantities: to $\mu^x r^y v^z$. The dimensions of the retarding force are MLT^{-2} (see page 29). Thus

$$MLT^{-2} = (ML^{-1}T^{-1})^x L^y (LT^{-1})^z,$$
$$MLT^{-2} = M^x L^{-x+y+z} T^{-x-z}.$$

Equating the indices of M, L and T

$$x = 1,$$
$$-1-z=-2, \quad \therefore z=1,$$
$$-1+y+1=1, \quad \therefore y=1.$$

Hence we get $$F \propto \mu r v,$$

and see that the retarding force varies as the viscosity of the air, as the radius of the drop and as its velocity of fall.

We can often use dimensional equations to suggest in this simple way the most interesting part of a relation, when we should have to exhaust the resources of mathematical civilization to complete the solution of the problem and get the constant. The full equation, for instance, which gives the retarding force on a small falling sphere is

$$F = 6\pi\mu r v,$$

but it needed the genius of Sir George Stokes to work it out.

As a rain drop falls, its velocity, and therefore the retarding force, increase, while the weight remains constant. Eventually, therefore, the retarding equals the accelerating force, and the drop will then fall with a constant speed called the limiting velocity.

Since the weight depends on the volume of the drop $\frac{4}{3}\pi r^3$, and the retarding force varies as the radius, the limiting velocity must vary as r^2, and becomes very small when the drop is small. Hence very small drops fall with negligible velocity, that is, they float about in the air as mist or cloud. If their size increases, as it tends to do because the surface of contact and therefore the total surface energy are thus diminished, the drops fall faster and eventually become rain.

CHAPTER III

HEAT AND ENERGY

Temperature

THE science of heat has developed from attempts to explain our sensations of heat and cold. We know what we mean when we say that a thing is hot or cold, and we can arrange a series of bodies in order of their hotness, thus getting a first rough scale of temperature.

But we soon find inconsistencies in this scale. If we put one hand in hot water and one in cold, and then transfer them both to water which is just warm, we find that this same water feels hot to the hand which has previously been chilled, and cold to the hand which has been heated. Now a scale which is not consistent with itself is useless for scientific purposes, and we must look for some other means of measurement.

Nearly every body expands when heated, and the earliest thermometer, invented before 1597 by Galileo, depended on the expansion of the air in a bulb, the stem of which dipped in water. But the readings of this instrument were affected by the pressure of the atmosphere, and it was replaced about 1612 by alcohol sealed in a closed glass bulb, and in 1714 by mercury in glass, a combination which was found to be better in many ways. Next it was discovered that, once ice is formed, water freezes at a constant temperature, and also that it boils at a constant temperature at a constant pressure of the atmosphere. Hence the freezing and boiling of water can be used as fixed points. Marks are made on the glass tube at the height where the mercury stands at these respective points, and the length between the marks divided into equal parts to indicate degrees of temperature.

Celsius and Linnaeus, by dividing this length into one hundred degrees, invented the Centigrade scale, in which the freezing point of water is 0° and the boiling point 100°. Fahrenheit made his scale by taking as zero the coldest tem-

perature which he could find recorded, and the temperature of the normal human body as 100°. The freezing point on this scale came out about 32° and the boiling point about 212°. At a later date these exact numbers were accepted as the fixed points, giving the freezing and boiling points of water under a pressure of 30 inches of mercury. The temperature of the average healthy human body is then 98°·4. The Fahrenheit scale is still employed for ordinary purposes in England, but the Centigrade scale is used everywhere for scientific work, and for all purposes on the Continent.

With these thermometers it will be seen that temperature is defined in terms of the expansion of mercury in glass. 1° C. is that temperature above the freezing point of water at which mercury in glass has expanded by the one hundredth part of its expansion between the normal freezing and boiling points.

Clearly this scale depends on the properties both of mercury and of glass, and we shall not get the same scale if we change the liquid or change the kind of glass we use. The fixed points will be the same, but other temperatures will differ. Hence a temperature scale founded upon the readings of a mercury thermometer is an arbitrary scale. Convenience is the only reason why we choose mercury as our liquid and glass to make our bulb.

Air, which, as we have seen, was used in the earliest rough thermometer, may also be employed in an accurate modern standard instrument, though hydrogen or nitrogen are in some respects better. The readings of gas thermometers depend on the pressure of the atmosphere, but, when this is allowed for, they have the great theoretical advantage of agreeing among themselves.

One way of treating the effects of changing pressure is to keep the volume constant throughout an experiment by proper arrangements, and use the change in pressure itself to measure temperature. Fig. 16 shows a rough form of constant volume air thermometer.

The bulb *B* is filled with dry air or other gas, and plunged into a mixture of ice and water. The tube *GH* is raised or lowered till the mercury in the tube at *F* stands at the level

of a fixed mark. The pressure on the air in the bulb is then clearly that of the atmosphere together with that due to the higher level of the mercury in *H* over *F*. If the barometer stand at 760 mm. of mercury, and *H* be 10 cms. higher than *F*, the total pressure on or of the air is 76 + 10 = 86 cms.

The bulb *B* is now plunged into boiling water, or, better, into the steam from boiling water. The air expands, but it can be brought back to its original volume by raising *H* and

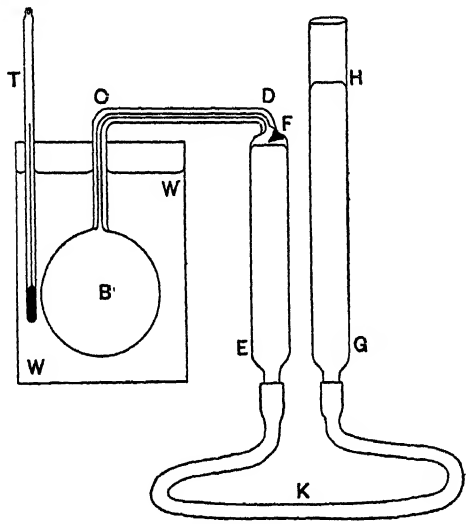


Fig. 16.

thus increasing the pressure and squeezing the air together. When the temperature is steady, *H* is adjusted till the mercury in *F* again stands at the fixed mark, so that the volume of the air is the same as at first.

Its pressure will now be higher, and will be measured by 76 cms. added to the new difference between the heights of *H* and *F* which will be found to be about 41·5 cms., a total of 117·5 cms. Between the freezing and boiling points of water, the pressure of air kept at constant volume has risen from 86 to 117·5, that is, from 1 to 1·366.

It is a very interesting fact that we get nearly this same

rise from 1 to 1·366 whatever gas, whether air, nitrogen, oxygen or hydrogen, we put in the bulb A. Within the limits of accuracy of such an experiment as this, these gases expand by the same fraction of their original volume. To this extent, a gas thermometer is less arbitrary than a liquid thermometer. Its readings depend, not on one fluid, but on a whole group of fluids—the less condensable gases—which all have the same thermometric properties.

Of course if we change the kind of solid used for the bulb, take porcelain for instance instead of glass, the result will be slightly different. But the expansion of gases, unlike that of liquids, is so much greater than the expansion of solids, that, except with much more delicate apparatus than we have described, the difference will be quite inappreciable.

Another kind of gas thermometer, instead of using the rise of pressure at constant volume, uses the increase of volume at constant pressure as the thermometric property. A bulb of glass or porcelain similar to that shown in Fig. 14 on page 44, is formed at the end of a fine tube, the weight of the whole being W_1. It is placed in the steam from boiling water, and the open end of the tube sealed up. The bulb is then removed, and plunged into water or mercury kept at the freezing point of water and the point of the tube opened. Liquid enters, and, when the level is the same inside as out, the pressure is that of the atmosphere as at first. The bulb is then weighed: the weight W_2 gives that of the bulb, together with the weight of liquid which fills the volume v by which the gas has expanded. Thus from $W_2 - W_1$ we can obtain v.

The bulb is then completely filled with the liquid and weighed again as W_3. V, the whole volume of the bulb, is calculated from $W_3 - W_1$. A volume $V - v$ at the freezing point, has become V at the boiling point. That is, a volume 1 has become $V/(V - v)$, or has increased by $v/(V - v)$.

It will be found that $1 + \dfrac{v}{V - v}$ is very nearly indeed 1·366, or the fractional increase of volume of a gas at constant pressure, is very nearly equal to the increase of pressure at constant volume.

This simplicity and regularity in these thermal properties

of gases is unlike anything we find in liquids or solids, and inevitably suggests to the mind that the gaseous state of matter is simpler in constitution than is the liquid state or the solid state. We shall have much to say later about this most interesting point.

Taking then either the constant pressure or the constant volume gas thermometer, we get a rise in either the volume or the pressure from 1 to 1·366 between the freezing and the boiling points of water. The fractional increase, or fundamental interval as it is called, is 0·366.

On the Centigrade scale, we have to insert 100 degrees in this interval, so that for each degree we shall get a rise of 0·00366 or 1/273 of the value at 0° C. At 100° C. the increase is 100/273, at 273° the pressure or volume is doubled, while, if we could cool the gas to − 273° C. and keep its properties unaltered, the pressure or the volume would vanish. This is the first indication we get of an absolute zero of temperature.

Now, of the ordinary gases, hydrogen is the farthest from its point of liquefaction, and, for very exact work, it is well to take a hydrogen thermometer as a standard of reference with which to compare other thermometers. Moreover, we shall see later on that the scale founded on the hydrogen thermometer very accurately, and the scales of other gas thermometers very nearly, agree with the so-called absolute or thermodynamic scale, which is based on the theory of thermal energy. An ordinary mercury thermometer can be compared with the hydrogen thermometer at a central station such as the National Physical Laboratory, and a table of corrections supplied to its purchaser. Temperatures as measured may thus be referred to the hydrogen scale without the tedious experiments necessary to use an accurate gas thermometer.

With this precaution, mercury thermometers made with special kinds of glass may now be used for accurate physical work, though for some purposes electrical thermometers have to be substituted.

Electrical thermometers are of three kinds (see also chapter IV):

(1) The electrical resistance of pure metals increases with

the temperature, and, curiously enough, to about the same extent as the volume or pressure of a gas. Hence a coil of platinum wire, wound on a mica frame and protected from gases by a porcelain sheath, may be used as a thermometer. Fig. 17 shows such a thermometer, and by its means a curve may be drawn in which the resistance of platinum is connected with the temperatures on the hydrogen scale.

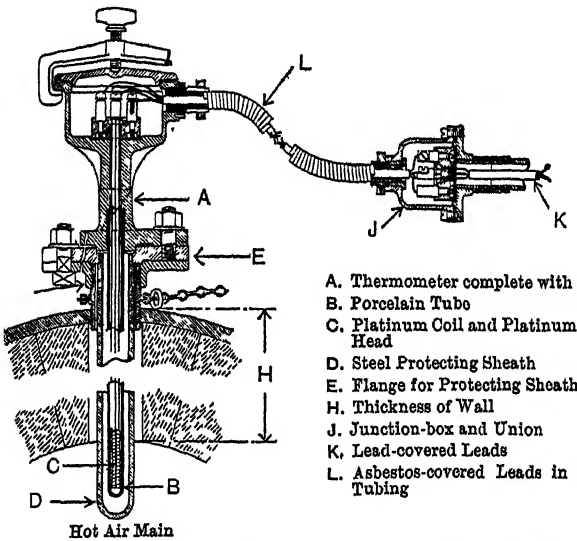


Fig. 17. Platinum resistance thermometer.

Platinum does not melt below 1800° C. and, although the glaze on a porcelain tube fuses about 1100°, higher temperatures may be reached by using a tube of silica. The platinum thermometer has therefore a very wide range, and, as the readings are obtained at a distance, may be used as a pyrometer for measuring the temperature of furnaces.

(2) If an electrical circuit be made of two different metals, and one junction be kept at a different temperature to the other, an electromotive force is set up and an electrical current may be obtained and measured. In some cases, the electromotive force increases regularly with the difference of temperature, and one of these thermo-couples, as they

are called, made of platinum as one metal and an alloy of platinum and rhodium as the other, is now much used in industrial processes. With a string recorder, in which the pointer of the electrical galvanometer is pressed each half-minute on to a revolving cylinder by a string moved by clock-work, a continuous record of temperatures may be obtained.

(3) For temperatures approaching and above the melting point of platinum, we have to depend on radiation. The total radiation from a hot body is proportional to the fourth power of the absolute temperature, that is, the temperature measured from the zero of the scale of the gas thermometer—273° C. below the freezing point of water. The radiation from a small opening into a furnace, or from a tube with its closed end inserted into the furnace, is focussed on to a thermo-couple. The deflection of a galvanometer can then be used as a measure of the intensity of the radiation, and therefore of the temperature of the furnace.

Quantity of Heat

When a hot body is put in touch with a cold one, the latter grows warmer, and we are accustomed to think of the process as the passage of something we call heat from the hotter to the colder body. Two contrasted views as to the nature of heat have been held from early days. Some have thought it to be an agitation of the particles of bodies, and others represented it as a subtle, invisible weightless fluid, permeating the pores of all substances.

At the end of the eighteenth century, when phlogiston was just vanishing from chemical philosophy, the theory of a caloric fluid was more clearly formulated than ever before, and started on a short but useful term of life. The advance in knowledge for which the time was ripe was a method of applying measurement to the phenomena of heat. For this practical development, the idea of a quantity of heat was necessary, and before the theory of energy had been established, it was easier to conceive of heat as a fluid than as a mode of motion.

In order to measure, we need a unit, and the unit of heat is best defined as the amount of heat required to raise the

temperature of a given mass of water through a given temperature interval. On the British system, we may take a pound of water and a degree Fahrenheit. On the C.G.S. system (see page 28) we choose a gram of water and a degree Centigrade. It becomes a matter of experiment whether the same amount of heat is needed at any degree, whether high or low, on the temperature scale. Observation shows that, while the amount is nearly the same, and may be taken as the same for rough work, differences do exist and must be brought into account for delicate experiments. It is necessary, then, to specify the particular degree Centigrade or Fahrenheit we select. The most convenient unit, and the one now generally adopted, is the amount of heat required to raise the temperature of one gram of water from 15° to 16° Centigrade.

The thermal capacity of a body is the number of thermal units needed to heat it one degree, and its specific heat is the number needed to heat it one degree compared with the number needed to heat an equal mass of water from 15° to 16° C.

The process of measuring a quantity of heat is known as calorimetry, and the piece of apparatus used is called a calorimeter. A simple calorimeter, shown in Fig. 18, consists of a copper vessel, suspended in air by threads inside another vessel, and containing a weighed mass of water. We

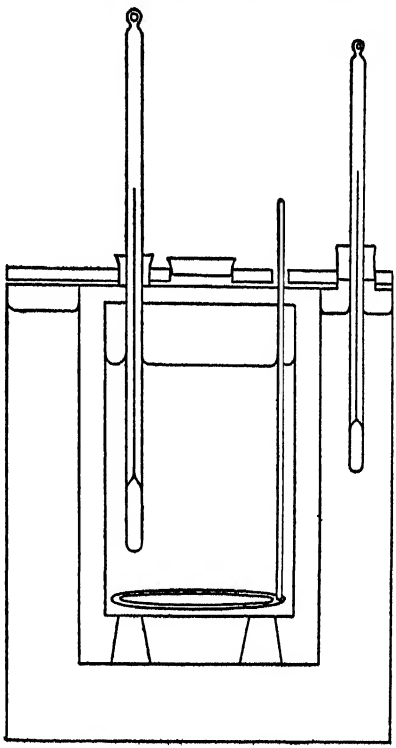


Fig. 18. Calorimeter.

will describe its use to determine the specific heat of a solid. A lump of the solid is weighed and hung up by a thread inside a tube surrounded by the steam from boiling water. When, after perhaps an hour, the temperature has become steady, the solid at a temperature T is dropped into the water in the calorimeter at a temperature t, the water is stirred and the final temperature taken. If the rise is great, correction must be made for heat lost during this process. The corrected temperature t', which the water would have reached had there been no such loss, gives the total rise in temperature of a known mass M of water and a vessel of known weight W and specific heat s'.

The total thermal capacity of the calorimeter and its contents is $M + Ws'$, and the total amount of heat absorbed is $(M + Ws')(t' - t)$. The amount of heat given out by the solid in cooling from the temperature of the steam is $(T - t')\,ms$, where s is the specific heat of the solid, and m its mass. These two quantities must be equal, and so we get

$$(T - t')\,ms = (M + Ws')(t' - t),$$

or

$$s = \frac{(M + Ws')(t' - t)}{(T - t')\,m}.$$

The only other calorimeter of importance is one devised by Joly and dependent on the condensation of steam. This instrument will be dealt with later.

Latent Heat

If a large lump of ice be broken and the fragments placed in a vessel with water, the water is cooled to the freezing point, and, when the whole is in equilibrium, it will remain at the temperature of the freezing point as long as any ice remains, or, as it is said, as long as the two phases, solid and liquid, are present in equilibrium.

If the vessel containing the ice and water be placed over a flame, heat is, of course, absorbed, and the ice rapidly melts. No rise in temperature occurs; hence the heat must be used up in melting the ice, and is said to become "latent." As soon as the ice is all melted, the temperature begins to rise and continues to do so until the water boils. Then, once more, it stays constant as long as water remains, as long, that is,

again, as two phases, liquid and vapour, are in equilibrium. Here another change of state occurs, from water to steam, and, while this is going on, the heat is used not in raising temperature, but in bringing about the change of state or phase. The heat does work in separating the water into the much more scattered molecules of steam, and becomes latent. When all the water is evaporated, the temperature of the steam begins to rise, and continues to do so till, at a very high temperature, the steam begins to dissociate chemically into hydrogen and oxygen.

The latent heat corresponding to these changes of state from solid to liquid and from liquid to vapour can be measured with the simple calorimeter described. The calorimeter is weighed first empty and then when containing water, the temperature of which is taken. A clean lump of ice is then dropped in, the water stirred till the ice is melted, and the corrected temperature recorded. The new weight of the vessel gives the weight of the ice added, and from the fall in temperature the latent heat of fusion, that is, the heat needed to melt one gram of ice, is calculated. It is found to be about 80 thermal units, the best modern values being about 79·8.

The latent heat of evaporation may be measured by passing a current of steam into the water of the calorimeter, recording the rise in temperature, and correcting for the loss in the process. Again, the final weight of the vessel gives the mass of steam passed in, and thus the latent heat, the heat needed to evaporate one gram of water, is found. Its value is about 538. Since the volume has been much increased in changing the water into steam, this figure includes the energy which has gone to produce expansion against the pressure of the atmosphere. Correcting for this, the internal latent heat is found to be about 498 thermal units.

The Conservation of Energy

The production of heat by friction is a matter of common observation. It bears a fairly obvious interpretation on the theory which regards heat as an agitation of the ultimate particles of bodies, but it is difficult to explain on the caloric fluid theory. The upholders of that theory were forced to

imagine that the thermal capacity of the filings or dust rubbed off from the substance by friction was much less than that of the substance in bulk, so that heat was squeezed out when the filings were made.

This explanation was disproved by Sir Humphry Davy, who rubbed together in a vacuum two lumps of ice, and melted them by the heat produced in the process. Now water was known to have a higher specific heat than ice, so the result of Davy's experiment was inconsistent with the fluid theory of heat. Nevertheless that theory still survived for nearly forty years, and received its death blow from another and different investigation.

For this result, the time was prepared by an accumulation of indirect evidence, which gradually became grouped under the name "correlation of forces." Force is here used as equivalent to what we now term energy; indeed, till Rankine defined the word energy as the power of doing work, the word force was used in two quite distinct senses to do duty for both force and energy, and much confusion of thought was the result: a science cannot develope satisfactorily till convenient fundamental quantities are conceived and accurately defined. The phrase "correlation of forces" stood for the connection between various forms of energy and the power of conversion of one into the others. Since mechanical work will produce heat, and heat can be reconverted into mechanical work in an engine, it became probable that heat was, like other "forces," a form of energy, an idea then expressed by calling heat "a mode of motion."

The problem was taken up about 1840 by James Prescott Joule, who set to work to measure accurately the heat developed by work of different kinds, mechanical and electrical.

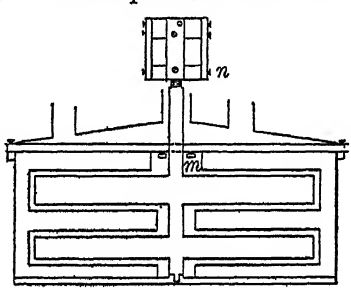


Fig. 19.

His first experiments were carried out on the heat due to mechanical friction. The water in a calorimeter (Fig. 19 and *h*

in Fig. 20) was churned by paddle wheels moving between fixed vanes, the paddle wheels being driven at first by falling weights, and in a later investigation by a hand wheel and cord (*d e* in Fig. 20). The amount of work done was measured from the couple exerted by the weights *k k*, and the quantity of heat developed was estimated from the rise in the temperature. Another experiment was made by churning mercury in an iron calorimeter, and yet another by heating a coil of wire immersed in water by an electric current of known strength.

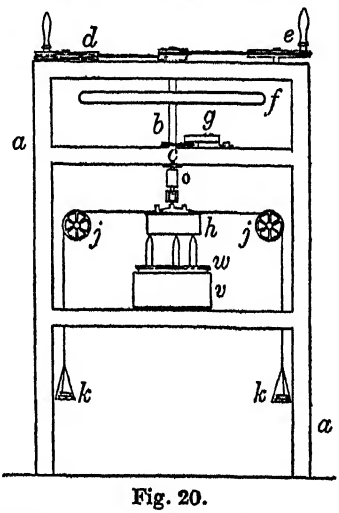


Fig. 20.

Joule found that, however the work was done and the heat developed, the expenditure of a given amount of work always produced the same quantity of heat—there was a fixed and definite proportion between them. Heat itself must be a form of energy, or a "mode of motion." That is Joule's permanent contribution to science. The definite figure he gave to express the proportion, namely first 772 and later 778 foot pounds of work to the thermal unit defined in terms of a pound of water and the degree Fahrenheit, has, of course, been superseded by more accurate modern work, but Joule's theoretical generalization still holds good. Heat is a "mode of motion" or a form of energy, and is convertible into energy of other forms.

The best modern experiments give a value of 783 in British units, and, in C.G.S. units, $4{\cdot}184 \times 10^7$ ergs as the dynamical equivalent of the gram-centigrade thermal unit (see page 62).

The vague statement of the correlation of forces was translated by Joule's work into the definite principle of the Conservation of Energy. The energy of an isolated system is constant in amount, it cannot be increased or diminished by any reactions between the parts of the system, though by

appropriate mechanism it may be made to undergo any of the changes of form of which energy is susceptible.

One of the earliest achievements of modern science was the proof of the conservation of matter through all the changes then known. Here we reach the corresponding result with regard to energy. It, too, is constant in amount and cannot be created or destroyed by any means at our command. Some people have gone so far as to say that the only two real things in the universe are matter and energy. That is a bold statement. It assumes that the mental model of nature which science gives at any one time is an accurate representation of reality, whereas history shows it may not even be permanent. Indeed, mass is now known to increase at very high velocities, and the recent study of radiation suggests that the product of energy and time, a quantity now called "action," is more important and fundamental than is energy itself.

However that may be, the concept we name energy is constant in all ordinary physical, chemical and biological changes, just as is the probably allied concept called matter. Indeed, it is probably because they are thus constant that they have emerged as fundamental quantities in physical science. Perhaps, then, the re-discovery of that constancy is not so very great an achievement.

The Specific Heats of Gases

The simple calorimeter already described does not lend itself to the measurement of the specific heat of a gas. Moreover, the apparatus designed for a gas must take into account the relation which holds between its pressure and its volume. Out of all the possible relations, the two most important, most interesting, and fortunately the most simple, are those two where the gas is kept either at constant pressure or at constant volume.

The measurement of the specific heat at constant volume presents great experimental difficulties, and has only been achieved directly of recent years by means of Joly's steam calorimeter. The measurement at constant pressure is comparatively simple and was carried out by Regnault about the year 1850.

The gas was compressed into the reservoir V, Fig. 21, and its pressure measured at the beginning and end of the experiment; this gave the mass of gas used. A steady current of the gas was allowed to flow by means of the stop-cock S through

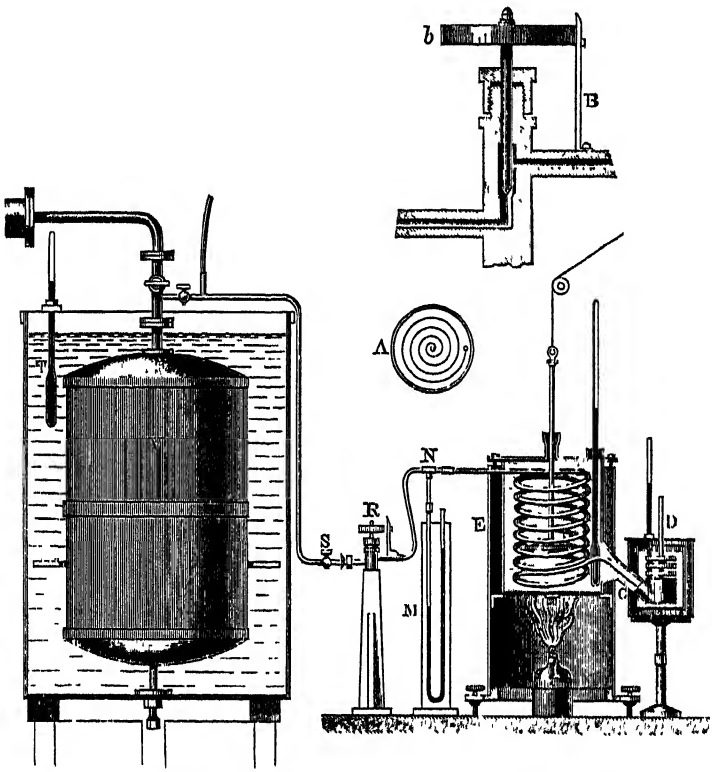


Fig. 21. Regnault's apparatus for specified heat of a gas.

the spiral tube immersed in the hot bath E at a temperature T. The gas then passed through another spiral in the water of the calorimeter C, which was heated from t to t'. The conditions are not the same as in the measurement of solids. There the whole of the solid remains in the water. Here the stream of gas passes on. The first part of the stream is cooled

to $t°$, the last part only to $t'°$. For moderate ranges of temperature we may average these two temperatures, and say that the gas has been cooled to the mean temperature of $\frac{1}{2}(t + t')$. If W be the total water equivalent of the calorimeter and its contents, we have

$$MC_p\left(T - \frac{t + t'}{2}\right) = W(t' - t)$$

whence C_p, the specific heat at constant pressure, may be determined.

The following figures give in the second column the specific heats (of one gram) of a few gaseous substances, and in the third column the specific heats per gram-molecule, that is, the chemical unit of mass, the molecular weight of the gas in grams. Since one gram-molecule of any gas which conforms to the simple gas laws occupies the same volume, these figures give the specific heats of equal volumes. It will be seen that, for the less condensable gases, these numbers are approximately the same, while for condensable gases, such as chlorine and carbon dioxide, they are higher.

Air ...	0·2374		Nitrogen ...	0·2438	6·83
Oxygen	0·2175	6·96	Chlorine ...	0·1210	8·57
Hydrogen	3·4090	6·82	Carbon dioxide	0·2020	8·89

In order to measure experimentally and directly the specific heat of a gas at constant volume, it must be confined in an inextensible vessel, the thermal capacity of which will be large compared with that of its gaseous contents. Hence the usual forms of calorimeter are unsuited for this purpose, and it was only when the more sensitive steam calorimeter was invented by Joly about 1888 that the experiment became possible.

A differential form of this instrument is illustrated in Fig. 22. Two similar globes of thin copper are suspended from the opposite arms of a balance by fine wires which pass through holes in the top of a closed chamber, into which steam can be admitted. Below the two globes are inverted cones, which serve to catch the water formed by the condensation of steam on the surface of the globes.

The globes are first exhausted by means of an air pump, and any difference in weight observed on the balance. Steam

is then admitted to the enclosing chamber, and, when the whole apparatus is raised to the temperature of the steam, the weights are again adjusted. The water caught in the cones is the water condensed from steam in the process of heating the globes from the original temperature of the room to the boiling point. It should be very nearly equal in weight on the two sides. If there is any difference it gives the difference in thermal capacity of the two globes.

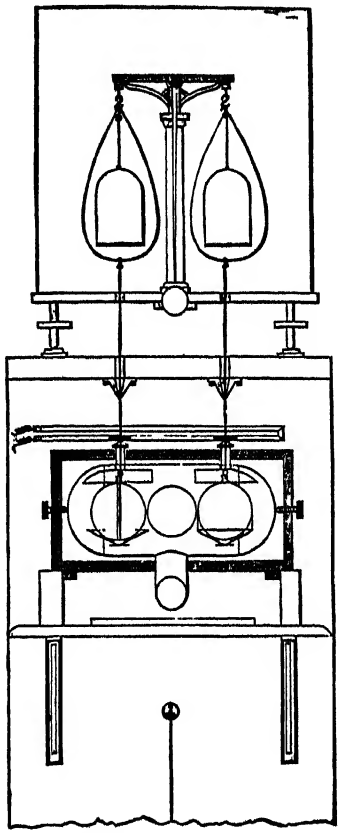


Fig. 22. Joly's differential steam calorimeter.

Into one globe air or the other gas to be investigated is then pumped to a pressure of from 7 to 20 atmospheres, and the experiment is repeated. Considerably more water is now collected on one side than on the other, and this excess gives by calculation the thermal capacity of the gas. If we know the volume of the globe and the pressure of the gas within, we can calculate the mass of the gas, and then, from the results of the experiment, deduce its specific heat at constant volume. Joly found a mean specific heat for air of 0·1721, the number for the lowest pressure he used, namely, 7·2 atmospheres, being 0·1684.

This is the only known way of measuring directly the specific heat at constant volume, but by several physical principles it is possible to determine the ratio of the specific heats at constant pressure and at constant volume. This ratio

is found to be about 1·40 or 1·41 for air. Taking the specific heat of air at constant pressure to be 0·2734, the specific heat at constant volume comes out 0·1692, a good agreement with Joly's result.

It is clear, then, that the specific heat of a gas at constant pressure is greater than that at constant volume. What is the reason for this difference?

If a gas be heated at constant pressure, it expands, and in doing so performs two kinds of work. Firstly, the molecules of gas are separated further from each other, so that, if there are forces of attraction between the molecules, work must be done in separating them. Secondly, an increase in volume carried out against an external pressure involves the performance of work. At constant volume neither of these kinds of work is done, and all the heat added goes to raise the temperature of the gas.

The first kind of work done at constant pressure, that needed to separate the molecules or the internal work, was examined experimentally in 1806 by Gay Lussac, and, independently, in 1844, by Joule. Fig. 23 shows Joule's apparatus. One of the two vessels *A* and *B* contains air at a pressure of several atmospheres, and the other is exhausted by means of an air pump. The vessels are placed side by side in the water of a calorimeter. When the temperature has become steady, the stop-cock between the vessels is opened, so that the compressed air expands and fills both vessels instead of one. Since the gas has expanded into a vacuum, no external work is done, the process being what is called free expansion. But, since the volume of the air is doubled, if there are attracting forces between the molecules, internal work is needed, which can only come from the heat energy of the gas. The gas will therefore cool. On the other hand, if the molecules repel each other, they will do work on expansion instead of absorbing it, and the gas will become

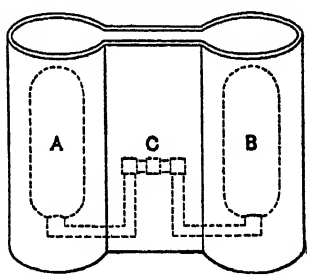


Fig. 23. Joule's apparatus for free expansion.

hotter. Joule found that when the water in the calorimeter was kept stirred, no appreciable change of temperature followed the opening of the stop-cock. The inter-molecular forces, if such there were, could not be detected by means of this apparatus.

A more sensitive form of apparatus was devised at a later date by Joule and William Thomson (afterwards Lord Kelvin) working together. An air pump, worked by an engine, forced a steady stream of gas through a coil of tubing immersed in cold water and then through a porous plug made of compressed silk and fixed in a boxwood cylinder. As the gas passes through the plug its pressure falls from that produced by the engine to that of the atmosphere. The gas therefore expands and does external work, but the following layers of gas are immediately compressed by the oncoming stream, the work needed being supplied by the engine. Hence the heat energy stored in the gas is not drawn upon for external work, and any change in temperature as the gas goes through the plug is due to internal work, work done on or by the inter-molecular forces, only. Thomson and Joule found slight cooling effects, indicating small forces of attraction, with most gases, but a still slighter heating effect with hydrogen, in which, therefore, there are minute forces of repulsion. These effects are proportional to the differences of pressure on the two sides of the plug, and the cooling increases as the temperature is lowered. Later experiments have shown that the heating effect with hydrogen, on the other hand, diminishes with falling temperature, and that at about $-80°$ C. it is reversed. At lower temperatures hydrogen is cooled like other gases.

The internal work of the less condensable gases having thus been shown to be small, it may be neglected except when great accuracy is needed. The difference between the specific heat of a gas at constant pressure and that at constant volume must therefore be almost entirely due to the external work done in expanding the gas against the pressure of the atmosphere, and can easily be calculated.

Let us suppose that a quantity of gas is confined in an engine-cylinder, the piston of which moves freely without

friction. Let the pressure be p and the volume of the gas v. If we heat the gas through a small temperature range dT, the volume will increase by a small amount dv, the piston rising through a small height dh. If the area of the piston be A, the total force on it is pA, and the work done on expansion $pA.dh$. But $A.dh$ is equal to dv, the change in volume, so that the work done is $p.dv$. The work done in any large change is the sum of all these small quantities, $\Sigma(p.dv)$, but, if we keep the pressure constant, this is the same as $p.\Sigma dv$, or $p(v_2 - v_1)$, where v_1 is the first and v_2 the final volume.

Now let us suppose that the gas consists of one gram of air, which will fill 772 c.c. at the standard temperature and pressure, 0° C., and 76 cms. of mercury, and that it be heated through 1° Centigrade. The work done will be (see page 46)

$$W = p(v_2 - v_1) = 76 \times 13{\cdot}6 \times 981 \times \tfrac{1}{273} \times 772$$
$$= 1{\cdot}014 \times 10^6 \times 2{\cdot}83 = 2{\cdot}87 \times 10^6 \text{ ergs.}$$

The thermal measurements we have described give this same difference in specific heats as $0{\cdot}2374 - 0{\cdot}1684 = 0{\cdot}0690$. These numbers are expressed in calories, so that comparison with the work calculated above should, if our theory and calculations are correct, give another value for the mechanical equivalent of the thermal unit agreeing with that deduced from the experiments of Joule and his followers. We have

$$J = \frac{2{\cdot}87 \times 10^6}{0{\cdot}069} = 4{\cdot}16 \times 10^7$$

a number which agrees extremely well with the entirely independent value got by direct experiment (page 64). It will be seen that this calculation of J from the specific heats depends on the experimental result that the internal work of air is small so that the external work only need be considered.

The Kinetic Theory of Gases

Joule's experiments were soon accepted as establishing the view that heat was the energy of motion of the molecules of bodies. Joule himself gave the first sketch of a detailed theory of the simplest case—that of the gas. He explained the pres-

sure by the impact of the molecules on the walls of the containing vessel, and temperature as the kinetic energy of translation of the molecules. On this theory he deduced many of the properties and relations of a gas as known experimentally.

Let us picture to ourselves a small cube, each side of which is one centimetre in length, filled with n molecules of a gas such as hydrogen. The molecules will be moving in all sorts of directions, but the net result will be somewhat the same as that produced by three streams of molecules each $\frac{1}{3}n$ in number moving backwards and forwards at right angles to each opposite pair of faces of the cube. When a molecule of mass m and velocity V strikes the wall, it rebounds with equal speed, so that the momentum suffers a change of $2mV$. The molecule traverses the space from wall to wall V times per second, and strikes one face $\frac{1}{2}V$ times per second. The rate of change of momentum is therefore $2mV \times \frac{1}{2}V = mV^2$, for one molecule, or $\frac{1}{3}nmV^2$ for all the molecules in one stream. But this rate of change of momentum measures the force on the face, that is, since its area is unity, the pressure exerted on it by the impact of the molecules is $p = \frac{1}{3}nmV^2$.

Now nm measures the total mass per unit volume, that is, the density d of the gas, so that we have $p = \frac{1}{3}dV^2$. The volume v of unit mass is $1/d$, so that $pv = \frac{1}{3}V^2$.

According to Joule's theory, the energy of a gas increases with the temperature, and temperature is given by the average kinetic energy of a molecule, or $T = \frac{1}{2}mV^2$.

Thus we get

$$pv = RT,$$

where R is a constant, an equation, which expresses both the experimental result known as Boyle's Law and also gives a definition of temperature on the gas scale.

If we have two gases, such as oxygen and hydrogen, at the same temperature and pressure, we get

$$\tfrac{1}{3}d_1V_1^2 = \tfrac{1}{3}d_2V_2^2 \text{ or } \frac{V_1}{V_2} = \sqrt{\frac{d_2}{d_1}}.$$

This explains what is called Graham's law of gaseous diffusion, which states as a result of observation that equal volumes of

two gases diffuse through a porous wall at rates that are inversely proportional to the square roots of their densities.

Once more the equality of temperature and pressure gives

$$\tfrac{1}{3}n_1m_1V_1^2 = \tfrac{1}{3}n_2m_2V_2^2 \quad \text{and} \quad \tfrac{1}{2}m_1V_1^2 = \tfrac{1}{2}m_2V_2^2,$$

and, dividing one equation by the other, we get $n_1 = n_2$ and thus confirm Avogadro's hypothesis. It is clear that Joule's theory is in conformity with the known experimental relations of gases.

Since in the equation

$$p = \tfrac{1}{3}\rho V^2,$$

the other quantities can be determined, we can calculate V, the average velocity of the molecules. At 0° C. the molecules of hydrogen on these figures move with an average velocity of 1840 metres per second, and those of oxygen at a rate of 461 metres per second. A more exact form of kinetic theory leads to results not differing much from those given by Joule's simple account. The more exact calculation indicates average molecular velocities at 0° C. of 1700 metres per second for hydrogen and 425 metres per second for oxygen.

For a substance which conforms to Boyle's law, it is easy to calculate the numerical value of the gas constant R. Indeed, its value is the same for all such substances, if the molecular weight in grams be taken. Thus, 2 grams of hydrogen occupy 22·32 litres or 22,320 cubic centimetres, when the pressure is that of the standard atmosphere, 760 mm. of mercury, and the temperature 0° C. or 273° on the gas scale, the freezing point of water. Hence

$$R = \frac{pv}{T} = \frac{76 \times 13{\cdot}6 \times 981 \times 22320}{273}$$

$$= 8{\cdot}14 \times 10^7 \text{ ergs per degree.}$$

It is sometimes convenient to express this value, which clearly represents the external work done in heating one gram-molecule of any gas through 1° C., in terms of thermal units or calories. To do so, we have to divide by J, the mechanical equivalent of heat, $4{\cdot}18 \times 10^7$. It will be seen that R is approximately equal to 2 calories per degree.

In monatomic gases such as argon, we may imagine all the heat absorbed to go to increase the kinetic energy of the

molecule as a whole, that is, to raise the temperature. Hence the specific heat per gram molecule at constant volume is

$$C_v = \tfrac{1}{2}\, nmV^2 v \times \frac{1}{T},$$

where v is the volume occupied by one gram-molecule.

But the pressure $p = \frac{1}{3}nmV^2$. Hence

$$C_v = \frac{3}{2}\frac{pv}{T} = \frac{3}{2}R = 3 \text{ calories per degree,}$$

since, as we saw above, R is about 2 calories per degree. An experimental determination on argon gave the specific heat at constant volume as 2·98.

In solids the molecules vibrate about a mean position, so that half the total energy is potential and half kinetic as in the case of the spring and weight on page 37. Hence the total energy is twice the kinetic energy, and the specific heat per gram molecule, which is called the atomic heat, in the case of a monatomic solid element should be twice the value for a monatomic gas, or 6 calories per degree.

Experiment shows that the atomic heats of metals range round this constant, e.g. sodium 6·4, copper 5·6, iron 5·9, silver 5·8. The values fall off at low temperatures, and even at ordinary temperatures some solid elements such as carbon have low atomic heats, though they approach the theoretical value at high temperatures. We may fairly conclude that solid metals have monatomic molecules.

Heat Engines and the Absolute Scale of Temperature

The theory of steam engines and other heat engines was put on a sound footing by the Frenchman Sadi Carnot, a son of the "Organizer of Victory" of the first Republic. Carnot pointed out the importance of simplifying the problem by imagining a theoretically perfect engine put through a complete cycle of operations. Just as we begin the study of falling bodies by neglecting the resistance of the air, so, in a simplified engine, we may suppose that there is no leakage of heat and no friction; the piston moves quite freely in either direction and its motion is reversible.

All experience shows that, in order to get steady work from an engine, heat must be supplied to the engine at a high temperature and the waste heat taken away by a condenser at a lower temperature. For instance, an engine cannot work at the expense of the unlimited store of heat in the sea or the earth; we have to buy expensive coal or other fuel to get a temperature higher than that of the surroundings.

Now the greater the difference of temperature between the source and the condenser, the better the engine works, and the less of the total heat supplied is rejected as waste. Hence, as Lord Kelvin saw, it is possible to use the ratio between the heat H supplied and the heat h coming out in an ideal engine, as a new means of defining temperature. The ratio of two temperatures T and t on this absolute or thermodynamic scale is defined as equal to the ratio between the heat absorbed and that given out by a perfect reversible engine working with a source and a condenser at the two temperatures. That is

$$\frac{T}{t} = \frac{H}{h}.$$

If the working substance in the cylinder of an ideal engine is a gas such as hydrogen, in which, as we have seen, the internal work of expansion is negligible, the heat put in must all go to do work at the high temperature, and the heat which comes out is due solely to work done at the low temperature. These two quantities of work may be shown to be proportional to the two temperatures on the scale of the gas thermometer, and these two temperatures must therefore correspond to the temperatures on the absolute scale, any slight divergence being calculable from the results of the porous plug experiment described on page 70. It follows that absolute temperatures may be measured by a hydrogen thermometer.

We define the efficiency of an engine as the ratio of the work done to the heat absorbed from the source, or the fraction of the heat taken in which the engine will turn into useful work. With Carnot's ideal engine, since there is no heat lost, the work done, W, is equivalent to the excess of heat taken in over that which is given out:

$$W = H - h,$$

and the efficiency $E = \frac{W}{H} = \frac{H-h}{H}$.

Hence, from what has been said above,

$$E = \frac{T-t}{T}.$$

This result shows that, even with a perfect engine, it is impossible to transform the whole of a given quantity of heat into work unless the condenser is at the absolute zero, − 273° C. At that point t is zero. Our equation gives $E = 1$, and shows that the efficiency is unity, or the whole heat is turned into work. At possible temperatures, the efficiency is much less, even for a perfect, ideal engine.

A high pressure steam engine may work at a pressure of 150 lb. to the square inch, which means that the water in the boiler is at a temperature of about 180° C. If the waste steam is emitted into the air, it will be at a temperature of 100° C. Adding on 273°, we get 453° and 373° as the temperatures of the source and the condenser on the absolute scale. Hence the maximum possible efficiency of such an engine is only $\frac{80}{453}$ or less than $\frac{1}{5}$, while of course friction, conduction of heat and other imperfections reduce the actual efficiency to a fraction considerably smaller.

One great advantage of the internal combustion engine, now so much used in motors, is the high temperature reached by the exploding gases in the cylinders. This makes the theoretical efficiency much higher than with steam engines, and the actual efficiency is higher also.

The Dissipation of Energy

Every real change which goes on in the world involves some leakage of energy by friction, conduction of heat, etc. Hence, in an isolated system, differences of temperature are always tending to decrease, and it becomes less and less easy to transform heat into work. In other words, the availability of the energy tends to become less, though, of course, the total amount of energy present is unchanged.

This conclusion, which is sometimes known as the principle of dissipation of energy, has been applied boldly to the whole

physical Universe, in which, it is argued, the store of energy must become gradually less available, till it exists altogether as uniformly distributed heat, and no further mechanical or electrical work can be obtained. This argument assumes firstly that the Universe may be treated as an isolated system, and secondly that no reconcentration of energy can occur. That such a reconcentration is at least conceivable was shown by Clerk Maxwell, who imagined a being able to discriminate between individual molecules. Such a being, which Maxwell called a daemon, could watch the different molecules of a gas. He or she could allow those moving faster than the average to pass a frictionless trap door, and thus concentrate energy on the farther side.

Nevertheless, the conception of availability is of supreme importance, and the theory of available energy plays a very great part in modern physics and perhaps even more in chemistry. Since the availability of energy tends to diminish, systems will be in equilibrium when their available energy is a minimum, and on this basis, as we shall see in chapter v, a complete theory of chemical and physical equilibrium has been built up.

Vapour Pressure

If a glass tube *AB* (Fig. 24) be filled with mercury and inverted with its open end in a trough of mercury *E*, it becomes a barometer. The mercury in the tube falls till it stands about 30 inches or 760 mm. above the level of the liquid in the trough. At this level the pressures inside and outside the tube must be equal. Inside the pressure is that of the column of mercury, and outside it is that of the atmosphere pressing on the surface of the liquid in the trough. Hence the height of the mercury in the tube over the surface of the mercury in the trough measures the pressure of the atmosphere.

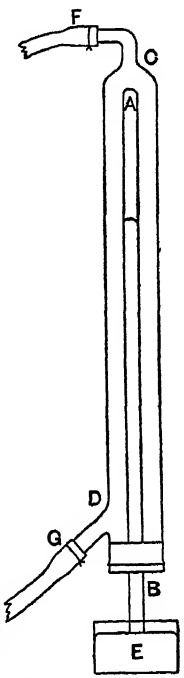


Fig. 24. Vapour pressure apparatus.

If a drop of water be placed below the open end of a barometer tube, it will rise through the denser liquid and float on the top of the mercury in the tube. Some or all of it will evaporate, and the pressure of the vapour so formed will depress the column of mercury by a height which measures the pressure of the vapour. If all the water has disappeared, another drop or two will serve to saturate the space, and leave a trace of water floating on the top of the mercury in the tube. It is found that further addition of water will not depress the column any more (except by the very small weight of the water itself). The system is now in equilibrium between the two phases of water and its vapour at the given temperature and pressure. If an attempt be made to alter the pressure, as, for instance, by inclining the tube, condensation of vapour follows, and the vertical height of the mercury in the tube recovers its former value.

The effect of temperature may be investigated experimentally by enclosing the tube in a water bath *CD* (Fig. 24). It is well to place by its side another barometer tube in which there is no water, for purposes of comparison. When the bath is heated, the level of the mercury in the tube with water falls, showing that the pressure of the vapour has increased. If it be cooled, the level rises to its original value. A definite height corresponds to each temperature.

Now let us approach the problem from another side. Water boiling in an open vessel is pushing out its vapour against the pressure of the atmosphere and doing a corresponding amount of mechanical work. If the pressure of the atmosphere falls, as it does if we climb a mountain, the amount of external work needed to evaporate the water becomes less, and therefore the water boils more easily, that is, at a lower temperature. The boiling point on a mountain is less than at sea level, and, indeed, the height of a mountain may thus be roughly estimated.

A less laborious method of experiment is to boil our water in a closed vessel connected with an air pump. The pressure of the artificial atmosphere over the water may then be easily reduced at will, and the boiling point corresponding to each temperature observed. A convenient apparatus for this

purpose is shown in Fig. 25. The steam from the boiler *B* is condensed in *CC*, and runs back as water into the boiler. The pressure of the air in the vessel *V* is adjusted by an air-pump connected with *A* and measured by the pressure gauge *M*. The boiling point corresponding to each pressure is observed on the thermometer *T*. By means of this apparatus, a curve may be drawn as shown in Fig. 26, giving the boiling points corresponding to different pressures for some way below the normal pressure of the atmosphere. Where this curve overlaps the former one, the two will be found to coincide—the two methods of investigation lead to the same results, and, by combining them, a complete curve extending from the freezing to the boiling point may be obtained. This curve is shown in Fig. 26.

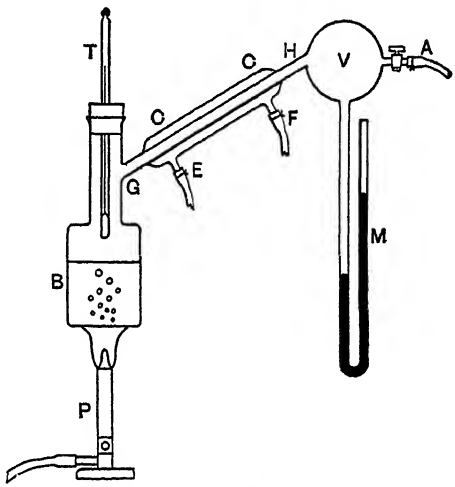


Fig. 25. Boiling point.

By carefully cooling the bath in Fig. 24, page 77, which must now be of some other liquid, say a salt solution, below the freezing point of water, the water in the tube may be cooled below its usual freezing point without the formation of ice, and the vapour pressure curve of undercooled water be recorded for a few degrees. By adding a crystal of ice and thus solidifying the undercooled water, the vapour pressure of ice may be measured. It will be found to be less than the vapour pressure of undercooled water at the same temperature.

At the other end of the scale, the vapour pressure of water above its normal boiling point may be measured by forcing air into the vessel of Fig. 25 instead of pumping it out. The water will now boil at higher temperatures, and the experi-

ment may be continued till the temperature is so high that the resulting pressure becomes dangerous.

Table of saturated vapour pressure of water.

(In millimetres of mercury.)

− 10° C.	2·160 mm.	120° C.	1,489 mm.
0	4·579	140	2,709
10	9·205	160	4,633
20	17·51	180	7,514
40	55·13	200	11,647
60	149·2	220	17,380
80	355·1	240	25,170
100	760·0	260	35,760

The vapour pressure of ice at −10° C. is 1·963 mm. and at 0° C. 4·579.

The experiments described for water can, of course, be performed with other liquids either more or less volatile than water. The same kind of results are obtained, and curves similar in general form may be drawn to express those results. In every case, it will be found that to a definite temperature corresponds a definite pressure. If we fix arbitrarily one of these two variables, the other is fixed automatically also. This fact may be expressed by saying that this system, made up of two phases, liquid and vapour, has one degree of freedom—we can freely choose one variable, but, when we have done so, equilibrium can only exist at a definite, single value of the other variable. To this point we shall return.

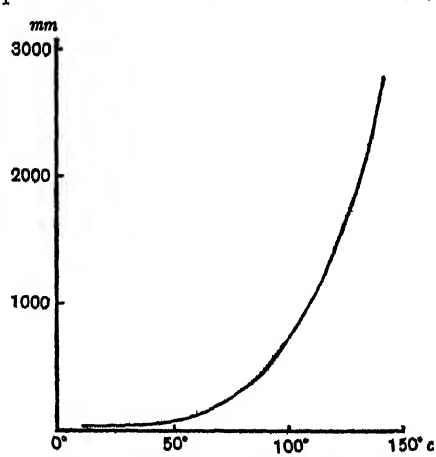


Fig. 26. Vapour pressure of water.

The Latent Heat Equation

The rate at which vapour pressure increases with temperature may be related with the latent heat and change of volume on evaporation by the thermodynamic principles we have already traced.

Let us imagine the cylinder of Carnot's ideal reversible engine to contain a liquid and its saturated vapour. When heat enters at the higher temperature T, evaporation takes place at constant pressure and is represented by the horizontal straight line AB in Fig. 27. The evaporation of one gram absorbs the latent heat L. Let condensation be carried out along the constant pressure line CD at a temperature only infinitesimally lower than T. Let us call this temperature $T - dT$. This will occur at a constant pressure just below the former pressure, say at $p - dp$, where dp is written likewise to denote a very small change in p.

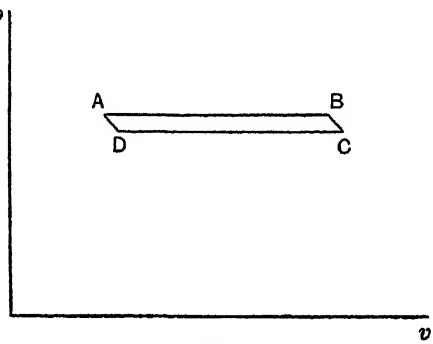


Fig. 27.

The work done in evaporation (see page 71) is $p\,(v_2 - v_1)$, and that absorbed in condensation $(p - dp)\,(v_2 - v_1)$, so that the work done by the engine during the cycle is $dp\,(v_2 - v_1)$. Hence by the equation on page 76 for the efficiency of such an engine

$$\frac{W}{H} = \frac{dp\,(v_2 - v_1)}{L} = \frac{dT}{T}$$

or

$$\frac{dp}{dT} = \frac{L}{T\,(v_2 - v_1)},$$

where dp/dT denotes the change in pressure divided by the change in temperature, that is, the change in pressure per degree Centigrade.

Now the latent heat L of evaporation of water at the standard boiling point, 373° C. absolute, is 538 thermal units,

or $538 \times 4{\cdot}18 \times 10^7 = 2{\cdot}25 \times 10^{10}$ mechanical units of energy. The change in volume $v_2 - v_1$ of one gram of water boiled into steam is 1649 c.c. Hence

$$\frac{dp}{dT} = \frac{2{\cdot}25 \times 10^{10}}{373 \times 1649} = 3{\cdot}66 \times 10^4 \text{ dynes per sq. cm. per degree}$$

$$= 27{\cdot}4 \text{ mm. of mercury per degree.}$$

This result may be verified experimentally by comparison with a table of vapour pressures giving more details than that on page 80. At 98° C. the equilibrium pressure is 707·3 mm. of mercury and at 102° C., 4° higher, it is 815·9 mm. This gives a change of $\frac{108{\cdot}6}{4} = 27{\cdot}2$ mm. of mercury per degree, a good agreement with the calculated number, and a satisfactory verification of the theory.

General Relations between the Liquid and the Gaseous State

Hitherto we have been dealing with substances which are liquid at ordinary temperatures. But, of course, similar relations hold for bodies which are ordinarily solids or gases. All solids melt or volatilize if heated to a high enough temperature, and all gases will condense if cooled and compressed. The true relations between the liquid and the gaseous state were first clearly elucidated by some experiments of Andrews on carbon dioxide or carbonic acid gas.

Andrews' apparatus is shown in Figs. 28, 29 and 30. Carbon dioxide was passed through the capillary tube *fba* for some hours to free it from air as far as possible. The end *a* was then sealed up, and a pellet of mercury *mn* introduced at the other end to isolate the gas. The tube was then securely fixed inside a copper tube with brass flanges (see Fig. 29), and the whole apparatus, up to the bottom of the pellet of mercury, filled with water. By means of a side tube (*ab* in Fig. 30) this water was put into connection with that in a similar tube, the glass of which contained dry air instead of carbon dioxide. The volume of this air was used as a measure of the pressure, which, owing to the side connection, was the same in both tubes. At the bottom of each tube was a screw plunger, which,

forced into the water, produced compression. The whole apparatus was then fixed in a water bath, so that the temperature was under control.

Fig. 28.

Fig. 29.

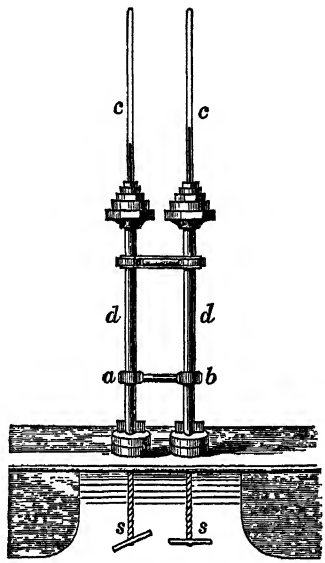


Fig. 30.

Andrews' apparatus.

Andrews recorded his results in the diagram reproduced in Fig. 31. When the temperature was kept at 13°·1, an increase in pressure began to liquefy the gas at about 49 atmospheres. At this value the pressure remained almost constant while the screws were worked in further and the whole gas liquefied. When no gas was left and the tube contained liquid only, a very small movement of the screw, that is a very slight decrease in volume, caused a very great increase in pressure; the curve at once goes up very steeply, almost parallel with the vertical axis of the diagram.

The process of liquefaction, when the gas is condensing to liquid and decreasing in volume at constant pressure, is represented on the diagram by a horizontal straight line. As successively higher temperatures are used, this horizontal part of the curve gets shorter and shorter, till, at 31°·1, it has almost vanished. At this temperature there is no visible liquefaction, though it is possible to see a layer of liquid formed at a temperature half a degree or so lower. Above this point no liquid can exist. Hence it seems that 31° C. is a critical temperature, above which no pressure, however great, can produce liquefaction.

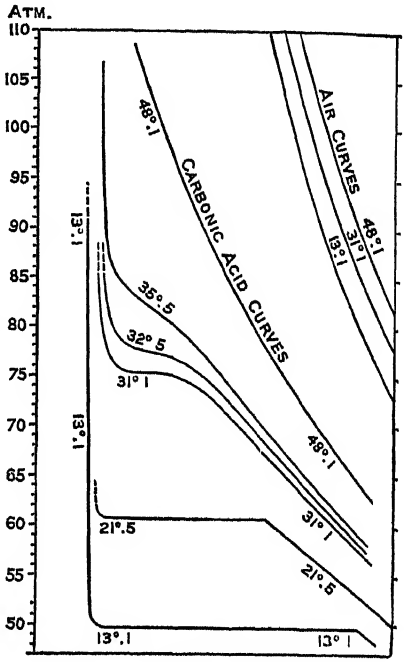


Fig. 31. Andrews' CO_2 diagram.

Andrews' experiments have been continued and extended by other physicists. Some of their results are given in the following table.

Substance	Critical temperature (° C.)	Pressure at the critical point (Atmospheres)
Helium	− 268	2·3
Hydrogen	− 234·5	20
Oxygen	− 118	50
Nitrogen	− 146	33
Air	− 140	39
Carbon dioxide	31·1	73
Ammonia	130	115
Chlorine	146	93·5
Ethyl alcohol	248	62·7
Water	365	194·6
Aniline	425·6	52·3

Returning to Andrews' diagram, we see that if carbon dioxide or any other substance be compressed above its critical point, then cooled, and finally allowed to expand, we can pass from a part of the diagram representing gas to another part representing liquid without touching the part where the curves are horizontal, that is, where visible liquefaction occurs with a definite separation into two phases. This passage of a single phase system kept homogeneous throughout from one phase to the other is often referred to as the "continuity of the gaseous and liquid states."

The Liquefaction of Gases

Andrews' work first clearly showed the conditions necessary for the liquefaction of a gas. If it be above its critical point, it is no use to increase the pressure. Hence the essence of the process of liquefying a resistant gas is cooling.

The final cooling is carried out by allowing the gas itself to do work with no access of heat. The energy needed for the work must then come from the heat energy of the gas.

On page 69 we saw that such work may be of two kinds—the external work done against the outside pressure, and internal work done against the attractive forces between the molecules. If a mass of gas be allowed to expand from a closed vessel into the atmosphere, both kinds of work are done. But, on the other hand, in the description of the porous plug experiment on page 70 we saw that, when gas is driven through a plug or an orifice by an engine and air-pump, it is the engine that does the external work and any cooling effect is due to internal work alone. At ordinary room temperature, this internal work is small, but it increases as the gas is cooled, and the advantage of a continuous process, in which a steady current of gas is driven along by an engine, is so great that it is better to lose the cooling effect of the external work and rely on the internal work alone.

The stream of gas is first cooled by some independent means. This is desirable with every gas and necessary with hydrogen, which is heated by expanding through an orifice if it be above $-80°$ C. By passing the stream through a spiral tube in a bath of boiling liquid air, it is cooled below this point.

Thus, whatever be the gas, when it expands through the orifice it is cooled. A simple laboratory form of apparatus is shown in Fig. 32. A spiral of copper pipe is wound round an exhausted tube, and fitted closely inside a vacuum vessel. Thus the colder gas is passed back over the coils of the tube through which the gas reaches the orifice *A*. The advancing stream of gas is thus made colder still, the porous plug effect is increased, and eventually the cooling is so intense that the gas is liquefied.

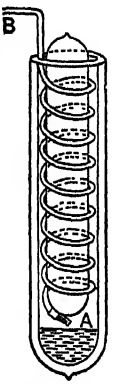


Fig. 32. Liquefaction apparatus.

In order to preserve liquefied gases, Sir James Dewar invented the vacuum vessel now so well known as the thermos flask. The space between two glass vessels is exhausted of air, and the opposite faces of the glass silvered. Metallic surfaces are bad radiators and absorbers of heat, and heat cannot pass by other means through a vacuum. Thus the rate at which heat leaks into a thermos flask is very small, and liquid air can be kept in an open flask for many hours very slowly boiling away. If the flask be corked up, the pressure and temperature will, of course, rise, and eventually an explosion follows.

Equations connecting Pressure, Volume and Temperature.

When a gas is above its critical point, as is carbon dioxide above 31° in Fig. 31 on page 84, the simple approximate relation known as Boyle's law holds good at constant temperature. The volume is found to be inversely proportional to the pressure, or

$$pv = \text{constant}.$$

If this result be expressed diagrammatically, we find that our observations lie on a rectangular hyperbola, if the diagram be plotted between p and v as axes (see Fig. 33), or a horizontal

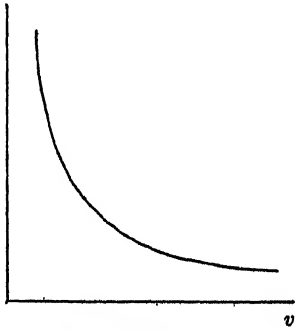


Fig. 33. Boyle's law curve.

straight line if it be plotted between pv as ordinate and p as abscissa.

Now let us consider the effect of a change of temperature. We have agreed to define temperature as proportional to the volume of a gas kept at constant pressure, or to the pressure of a gas kept at constant volume. And, as we have explained, temperature as thus defined is nearly in accord with the true absolute scale, founded on the principle of energy. Our definition of temperature on the gas scale, shows that the product pv is proportional to the temperature, or is equal to the temperature multiplied by a constant. Thus

$$pv = RT.$$

It is better to use this notation for a definite mass of gas, and it is most convenient to take the molecular weight of the gas (see pages 45, 67) in grams, or the gram-molecule, as the chemical unit of mass. Then, if we have m gram-molecules,

$$pv = mRT.$$

If we define temperature in terms of the expansion of hydrogen, other gas scales such as those of the air or the nitrogen thermometers will be nearly the same, and for them, too, the equation stands. But when we pass to substances like carbon dioxide, or indeed to any gas near its point of liquefaction, Andrews' curves show that more complicated relations must hold good. Many equations have been proposed to express these relations, the most famous, and in some ways the best, of which is that devised by the Dutch physicist Van der Waals.

A gas consists of molecules widely separated from each other, and, when it is compressed, it is natural to suppose that the volume by which it shrinks is the volume not occupied by the substance of the molecules themselves, or $v - b$, where b, the co-volume, is a correction for the latter, or to put it in another way, the molecules are forced nearer together but not themselves compressed.

Applying this correction, the equation for one gram-molecule becomes

$$p(v - b) = RT.$$

Next we have to consider any possible forces between the molecules. If such there be, it is clear that they will be proportional to the square of the concentration of the gas, since doubling the number of molecules will double them on each side of an imaginary plane, and thus double both the number attracting and the number attracted. Inter-molecular forces will therefore cause an extra pressure which may be written a/v^2, since v, the volume of unit mass, is inversely proportional to the density or the concentration. Correcting p for this added pressure, we have finally

$$\left(p + \frac{a}{v^2}\right)(v - b) = RT$$

or

$$p = \frac{RT}{(v - b)} - \frac{a}{v^2}$$

in either of which forms Van der Waals' equation may be expressed.

This equation is found to agree fairly well with the relations between temperature, pressure and volume for gases which depart somewhat from Boyle's law, and gives a rough approximation to the truth for substances of simple chemical constitution when the variation from Boyle's law is greater. The equation cannot express the sudden change from a curve to a horizontal straight line which we find in, for example, Andrews' curve for carbon dioxide. When plotted between p and v for axes, with appropriate values for the constants, the equation gives the observed results for the system when wholly liquid or wholly gas, and, instead of the horizontal lines in the figure, curves like a letter S lying on its side. Such a curve may be taken to represent the state of the substance in unstable equilibrium, when the vapour is supersaturated or the liquid overheated. Experiment can thus advance a little way along the S shaped line from each of its ends.

In Van der Waals' equation there are three constants. Hence it needs three determinations of temperature, pressure and volume to calculate the values of those constants. When these three determinations are made and the constants calculated, the equation may be used to estimate the tempera-

ture, pressure or volume at any point when the other two of these variables are fixed and known.

Andrews' diagram for carbon dioxide shows that the horizontal part of the curves vanishes at the critical point. Here the three points where the *S* shaped curve cuts the horizontal line coalesce, and the horizontal part of the curve vanishes. Those who like a simple mathematical exercise may throw the equation into the form of a cubic equation in v, and determine the critical constants by equating together the three roots of the cubic equation.

CHAPTER IV

ELECTRICITY AND MAGNETISM

Introduction

HITHERTO we have dealt with the phenomena of every-day life—with facts that cry aloud for explanation. No one could live, at any age of the world's history, without being faced by practical problems which depend on weight, on the differences between earth, water and air, on the changes from heat to cold, on the kindling of a fire and the danger of flame.

But, except for an occasional thunderstorm, all the mighty electric manifestations now so familiar were hidden from mankind till the curiosity of the philosopher discovered them and investigated their laws, and the practical genius of the scientific engineer saw and developed their technical application.

Electrostatics

One apparently insignificant fact known to the Greeks—the seed from which this huge tree of knowledge grew—was the attraction exerted on light bodies by amber when rubbed.

Dr Gilbert (1540–1603), physician to Queen Elizabeth, examined this force by means of a light needle balanced on a point, and invented the name electricity from ἤλεκτρον, amber. He discovered new electric substances, which became active when rubbed, and made them known by his writings. Stephen Gray in 1729 distinguished conductors from insulators. Dufay, working in France, discovered that there were two kinds of electricity, and, since one kind neutralized the other, Sir William Watson and Benjamin Franklin called them positive and negative. Similar kinds of electric charge repel each other and unlike kinds attract. Cavendish investigated the forces developed, and he and Coulomb independently found—as Newton had found before them with the hypothetical force of gravity—that the force varied inversely as the square of the distance between the two attracting particles.

Michael Faraday (1791–1867), working in London at the Royal Institution, showed that electricity could be treated as a quantity—that charges could be added or subtracted, and their forces changed in like proportion. He also examined the influence of the non-conducting or insulating medium—air, or whatever it may be, between the electric charges. The complete expression for the force F between two electrified particles may be written as

$$F = \frac{e_1 e_2}{k r^2}, \qquad \text{.....(1)}$$

where e_1 and e_2 are the electric charges on the particles, r the distance between them, and k a dielectric constant expressing the properties of the insulating medium or dielectric. The constant is least for a vacuum and nearly the same for air. This value is taken as unity. For ebonite it is about 2·8, for paraffin wax 2·2, and for glass it varies from 5 to 10.

From equation (1) we may get a convenient unit of electric charge. Unit quantity of electricity is defined as that quantity which, when placed at unit distance, say 1 cm., from an equal similar charge in air, repels it with a unit force of one dyne.

If a metal ball fixed on an insulating glass or ebonite stand be given a positive electric charge, and put near an insulated cylinder, the nearer end of the cylinder will acquire a negative charge and the further end a positive one. This process is called induction. If the cylinder be then touched with a finger for a moment, the positive induced charge is repelled to earth. The negative induced charge is left, but, as long as the original positive charge on the ball is near, the attraction keeps the opposite charges bound to each other. If the ball be removed, the negative charge on the cylinder is set free and distributes itself over the cylinder.

Larger quantities of electricity may be obtained either by frictional electric machines in which glass plates turned on an axis are rubbed between pads as they turn, or by induction machines, in which a series of induced charges are developed and removed.

In examining electric phenomena, some form of measuring instrument is needed. The simplest is the gold leaf electroscope. In its early form the electroscope consists of a glass

case containing two gold leaves hanging in touch with each other from a wire. When the wire is electrified, the gold leaves repel each other and stand out at an angle. The modern electroscope has only one leaf *C*, attached to a metal plate *D* (Fig. 34) from which the leaf is repelled when electrified. The gold leaf moves in front of a scale, the readings of which give its degree of electrification.

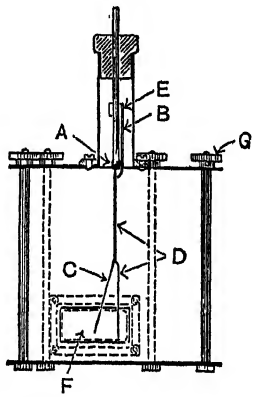


Fig. 34. Gold leaf electroscope.

The whole of an electric charge resides on the surface of a conductor. Just as we have to deal both with the volume of a vessel containing gas, and the pressure of the gas in it, so we have to consider the electric capacity of a conductor and the pressure or potential of the electricity on its surface.

The potential may be defined as the work done against the electric forces in adding unit quantity of electricity to the charge. If the capacity *C* is large, the system will absorb a large charge or quantity of electricity *e* before the potential *V* rises to a given value. The relation clearly is $C = e/V$, or the capacity of a system is measured by the charge required to raise its potential by unity. The capacity of a conductor may be increased by bringing near it another uninsulated conductor. The charge on the first will attract or induce a charge of the opposite kind on the second, and the attraction between the charges will reduce their potentials, just as an absorbent surface would form a film of gas and reduce its pressure in a vessel. Hence two metal plates, separated by a thin stratum of insulating material, will have a very large electric capacity. This arrangement is called a condenser. It is clear that, if the insulator have a high dielectric constant, the forces will be diminished and the capacity increased. Hence the specific inductive capacity of an insulator is the same thing as its dielectric constant.

These electric charges at rest gave rise to the important science of electrostatics, which was largely developed by

mathematical analysis. Though not altogether without practical interest, that science is chiefly, perhaps, of theoretical importance.

Magnetism

The loadstone, a magnetic oxide of iron, found in and named from Magnesia, was known to the Greeks, and the fact that, when rubbed on a steel needle, that needle tends to set north and south is said to have led the Chinese to the invention of the mariner's compass in very early times. It first appeared in European literature about the twelfth century of the Christian era.

Soft iron is only a magnet while acted on by the loadstone, or subjected to a magnetic force in some other way. But hard steel, when once magnetized, keeps its properties.

A steel bar magnet, suspended horizontally by a fibre of silk, sets nearly north and south. If a second bar magnet be brought near, it is easy to show (1) that the forces seem most intensive at points (called poles) near the ends, and (2) that the north-seeking end of one magnet repels the similar pole of the other and attracts the south-seeking pole. Magnetism, like electricity, is of two kinds or two signs. Magnetic poles of the same kind repel and unlike poles attract each other.

Again, as with electric charges, the force between two magnetic poles is found to vary inversely as the square of the distance between them, and to depend on the medium between the poles. The force may be written as

$$F = \frac{m_1 m_2}{\mu r^2}, \qquad \text{.....(2)}$$

where $m_1 m_2$ may be called the strengths of the magnetic poles, r denotes the distance separating them, and μ a quality of the medium called its magnetic permeability. It will be seen that μ corresponds to k in electrostatics, though the equivalence is not complete: k is a constant, but μ changes to some extent with the intensity of the magnetic force.

A pole has unit strength when it repels an equal similar pole at unit distance with unit force in air. The magnetic force at a point is defined as the mechanical force on unit pole placed at the point.

The distribution of force round a magnet is best seen by placing it beneath a card, and scattering iron filings over the card. Each filing becomes a little magnet, and they cling together, showing by their form what are called lines of force (Fig. 35).

The state of the magnetic field may be represented by drawing one line of force through each square centimetre for each unit of magnetic force in the strength of the magnetic field at that point.

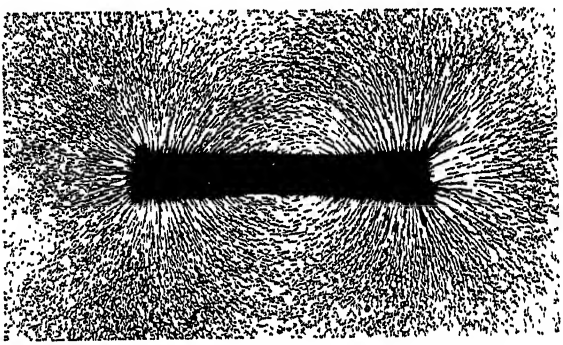


Fig. 35. Lines of magnetic force in iron filings.

A magnet will induce magnetism on a bar of iron placed near it, the north pole of the magnet inducing a south pole on the nearer end of the iron and an equal north pole on the further end. If we regard the lines of force as starting from a north-seeking pole and ending on a south-seeking pole, induction may be regarded as the appearance of south-seeking magnetism where lines of force enter iron and of north-seeking magnetism where they leave it.

The induced poles will themselves give rise to lines of force, and, since in the iron these secondary lines will run the opposite way to the primary lines, the induced poles produce a demagnetizing effect, greater the larger the cross section and the shorter the length of the iron. On the other hand, inside the substance of the iron the magnetism reinforces the primary magnetizing force. Consider a narrow crevasse of air

cut inside the iron as in Fig. 36. Where the lines of force leave iron and enter air, they induce a north pole and where they enter iron again a south pole. Hence in the crevasse the secondary lines run the same way as the primary force. Experiment shows that this secondary effect is very large, so that the ratio of the total magnetic force in iron, which is called the magnetic induction B, to the force H in air, a ratio which gives μ the magnetic permeability ($\mu = B/H$), may be as much as 3700 in soft wrought iron.

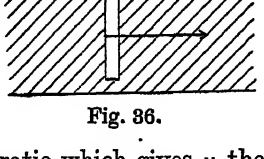


Fig. 36.

The fact that a compass needle sets north and south shows that the earth itself is a huge magnet. Its magnetic poles are not quite in the same places as the geographical poles, and move slowly as the years pass. The angle between the direction of the compass and true north is called the declination.

If a compass needle placed north and south be suspended so as to move freely in a vertical plane about a horizontal axis, its north-seeking pole will, in the northern hemisphere, dip towards the ground. The angle which the axis of the needle, when lying along the direction of the resultant magnetic force, makes with the horizontal is called the dip or inclination. This angle also varies slowly.

The following table shows the secular variations of the declination and the dip at Greenwich Observatory.

Date	Declination	Dip
1576		71° 50′
1580	11° 15′ E.	
1600		72° 0′
1622	6° 0′ E.	
1657	0° 0′	
1672	2° 30′ W.	
1676		73° 30′
1723	14° 17′ W.	74° 42′
1773	21° 9′ W.	72° 19′
1787	23° 19′ W.	72° 8′
1802	24° 6′ W.	70° 36′
1820	24° 34½′ W.	70° 3′
1860	21° 39′ 51″ W.	68° 19′
1893	17° 27′ 0″ W.	67° 30′
1900	16° 52′ 40″ W.	

If these two angles, the declination and the dip, be determined, and a measurement be also made of the horizontal component of the earth's magnetic force, a complete value of the earth's field at the point of observation is obtained.

If a short bar magnet be brought near a small compass needle, the needle will be deflected. If the resultant magnetic force of the bar magnet be directed east and west, the magnetic force may be calculated from the angle of deflection of the compass needle from the north and south line.

Magnetic force or magnetic field strength may be defined as the mechanical force on a pole of unit strength, and in Fig. 37 the magnetic force at P is due to both the poles of the bar AB, of strengths m and $-m$. Let l be the distance OA or OB from either pole to the centre of the bar and r the distance from the centre of the bar to a small compass needle at P. The total magnetic force at P is then

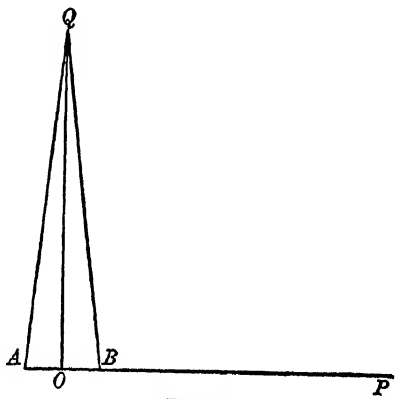


Fig. 37.

$$H = \frac{m}{(r-l)^2} - \frac{m}{(r+l)^2}$$

$$= m\left\{\frac{r^2 + 2lr + l^2 - (r^2 - 2lr + l^2)}{(r-l)^2\,(r+l)^2}\right\}.$$

Since the bar is short, l is small compared with r, and terms involving l^2 may be neglected. Hence

$$H = \frac{4mlr}{r^4} = \frac{4ml}{r^3} = \frac{2M}{r^3}. \qquad \text{.....(3)}$$

The quantity $2ml$ constantly recurs in these problems. It is written as M, and called the magnetic moment of the bar magnet. It will be seen that if the magnetic force of a pole varies inversely as the square of the distance, the total magnetic force due to a short bar magnet

varies inversely as the cube of the distance from its centre. This is easily verified, and forms one way of confirming the accuracy of the inverse square law for isolated poles.

The compass needle in Fig. 38 sets in the earth's magnetic field, the horizontal component of which we will write as **H**. Let it be deflected by the bar magnet through an angle θ or MCM'. The mechanical force to the north acting on M is $\mathbf{H}m$, and the component of this at right angles to the needle is

$$\mathbf{H}m\frac{MN}{CN} \text{ or } \mathbf{H}m\sin\theta,$$

and the moment of this force tending to turn the needle (see p. 36) is $\mathbf{H}ml\,MN/CN$. There will be an equal force and moment acting on the other pole, so the total moment or the couple is

$$2\mathbf{H}ml\frac{MN}{CN} \text{ or } \mathbf{H}M\frac{MN}{CN}.$$

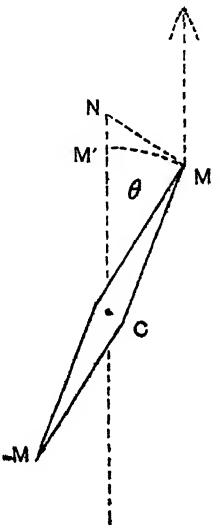


Fig. 38.

Let a magnetic force be applied at right angles to the earth's field—that is, east and west. The couple on the needle is

$$HM\frac{CM}{CN} \text{ or } HM\cos\theta.$$

When the needle is at rest these couples must be equal, and

$$HM\frac{CM}{CN} = \mathbf{H}M\frac{MN}{CN},$$

or

$$H = \mathbf{H}\frac{MN}{CM} = \mathbf{H}\tan\theta. \qquad \ldots\ldots(4)$$

Now $\frac{MN}{CM}$ may be measured experimentally, or we can find its value by tables if we measure the angle θ, for $\frac{MN}{CM}$ is $\tan\theta$. Thus, by observing the deflection of a compass needle acted on by an unknown magnetic force at right angles to the earth's field, we can find the strength of that unknown force H as compared with the horizontal force **H** of the earth's field.

The earth's horizontal force may be determined by deflecting

a compass needle with a magnet of moment M as in Fig. 37 on p. 96, and calculating the relation between **H** and M from equations 3 and 4. Then the same magnet is hung horizontally by a silk fibre. It will set in the earth's field, and if it be deflected and allowed to oscillate about its position of equilibrium, it will perform isochronous vibrations, the period of which gives another relation between **H** and M and thus enables us to calculate the value of each.

Current Electricity

An electric charge from a frictional or induction electric machine flowing along a conductor gives a current of electricity, but the quantity is too small for convenient investigation. It was only when Galvani and Volta discovered a new set of facts that large currents were obtained. In 1800 Volta invented the galvanic or voltaic cell, which, in its simplest form, consists of a plate of zinc and a plate of copper dipped in dilute sulphuric acid and connected outside the liquid by copper wires. A current, soon identified with electricity, was found to flow in a circuit through the liquid and back through the wire, the direction of the so-called positive electricity being from zinc to copper in the acid and from copper to zinc outside.

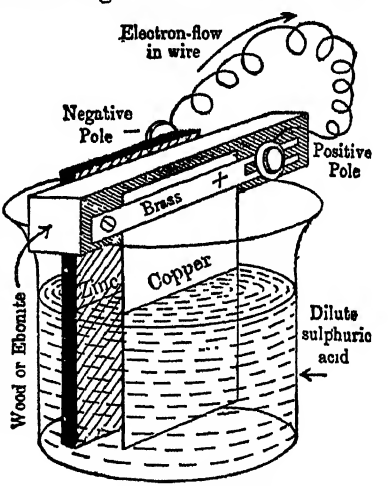


Fig. 39. Volta's Cell.

The current from this cell falls off rapidly, owing to the formation of a film of hydrogen on the copper plate. This action is called polarization, and may be prevented in several ways—for example, by placing the copper plate in a solution of copper sulphate, separated by a porous pot from the liquid round the zinc, which may be a solution of acid or of

zinc sulphate. This cell, invented by Daniell, may be described as

$$Zn \mid ZnSO_4.Aq \mid CuSO_4.Aq \mid Cn.$$

In it, zinc goes into solution and copper is deposited on the copper plate instead of hydrogen, and thus polarization is prevented. In other cells, the copper is replaced by platinum or carbon immersed in an oxidizing liquid such as strong nitric acid or a solution of potassium bichromate.

Another more practically useful cell is the lead accumulator, two porous lead plates in dilute sulphuric acid. This is charged by passing a current through it. A current may then be drawn off in the opposite direction. The main chemical action seems to be represented by

$$PbO_2 + Pb + 2H_2SO_4 = 2PbSO_4 + 2H_2O,$$

an equation, which, read from left to right, describes the discharge, and from right to left the charging of the cell.

The currents thus obtained produce marked chemical effects. When passed through salt solutions, they deposit metals at one end and evolve acids at the other. By their means, Sir Humphry Davy in 1806 decomposed the alkalis caustic potash and caustic soda, which were then thought to be elements, and liberated the metals potassium and sodium. The process of electroplating with silver, and the electrolytic method of purifying copper, are other illustrations of electrochemical action. We shall return to this subject below.

The Magnetic Force of a Current

If a wire carrying an electric current be held above and parallel to a magnetic compass needle, the needle is deflected from its usual north and south direction. Hence it follows that electric currents produce magnetic forces. A coil of wire suspended on a pivot and carrying a current sets in the earth's field and behaves at all points outside itself just like a magnet. Therefore at outside points a current is equivalent to a magnet of similar form and corresponding strength.

The wire and compass needle described above make a simple form of galvanometer, or instrument for measuring galvanic (i.e. electric) currents. The effect can be multiplied by winding the wire into a coil. If the circle of the coil be

large, the magnetic force near the centre is uniform over a considerable space, and a short compass needle suspended at the centre of a coil of which the plane lies north and south is deflected just as is the needle in Fig. 38 on p. 97. This may be used to define the strength of an electric current. The magnetic force is proportional to the length of the wire or $2\pi rn$, where n is the number of turns of wire in the coil, and inversely proportional to the square of the distance r of the wire from the centre. Hence, the current, c, being taken as proportional to the magnetic force, H, it produces, we may write from equation (4) on page 97,

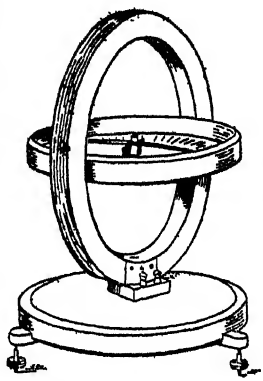


Fig. 40. Tangent galvanometer.

$$H = \frac{2\pi rnc}{r^2} = \frac{2\pi nc}{r} = \mathrm{H} \tan \theta,$$

where H is the horizontal component of the earth's magnetic force, and θ the angle of deflection of the needle. Therefore

$$c = \frac{\mathrm{H}r \tan \theta}{2\pi n}.$$

The instrument is called a tangent galvanometer, and by its means an electric current may be measured in C.G.S. units. A unit current, as thus defined in terms of its magnetic force, when passed through the solution of a silver salt, deposits 0·01118 gram of silver in one second. This current is rather large for practical convenience, and the tenth part of it—the current which will deposit 0·001118 gram of silver in one second—is defined as the practical unit of current and called the ampere.

Very often a galvanometer is used, not to measure a current in C.G.S. units, but merely to detect its presence and its direction. The sensitiveness may then be increased by making the diameter of the coil small, so that it closely surrounds the needle, and by decreasing with magnets the strength of the earth's field in which the magnet hangs.

Since a current is equivalent to a magnet, a moving-coil galvanometer may be made by suspending or pivoting a coil of wire between the poles of a horse-shoe magnet (see Fig. 41). When a current is passed through the coil it is deflected. As it is independent of the external magnetic field, this form of instrument is much used for practical purposes.

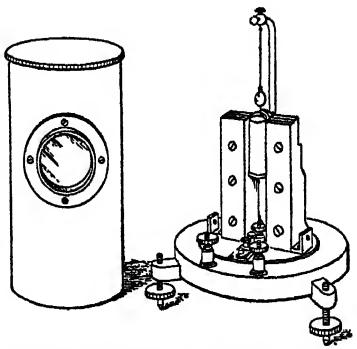


Fig. 41. Moving coil galvanometer.

It is obvious that the magnetic force of a current may be exerted at a distance from the battery producing it. Hence arose the electric telegraph, by which signals may be sent over line wires and through submarine cables.

Difference of Potential or Electromotive Force

When an electric current flows along a conductor, work is done, and the amount of work done between two points by unit current flowing for one second is called the difference of electric potential or the electromotive force (abbreviated as E.M.F.) between those points. Unit E.M.F. exists when unit work is done by unit current flowing for unit time. On the C.G.S. system this gives a unit too small for practical use, and a multiple of it, one hundred million or 10^8 times as great, is taken as a practical unit and called a volt.

A difference of potential is set up by the voltaic cell or other source of current. In a cell, the origin of the E.M.F. is the energy liberated by the chemical reactions, and its value may be calculated by thermodynamics from the heat equivalents of those reactions, which measure their chemical intensity, combined with the reversible heat effects at the junctions (p. 104).

The difference of potential between the terminals of a cell may be demonstrated and measured by a gold leaf electroscope or other form of electrometer. If two cells be connected in series, that is with the copper of one cell joined to the zinc

of the next, the E.M.F. of the combination is the sum of the single E.M.F.'s. On the other hand, two cells connected in parallel, with copper joined to copper and zinc to zinc, are equivalent to a single cell of larger plate-area, and give the same E.M.F. as a single cell, though, when the circuit is completed, the current this gives may be greater owing to their smaller resistance.

Just as a difference in pressure in a fluid results in a flow of the fluid from a place of high to a place of low pressure when a channel is open, so a difference of electric potential may be regarded as causing an electric current when the circuit is completed.

Ohm's Law and Electric Resistance

If the current be measured by a galvanometer and the E.M.F. be measured (for instance, by connecting more and more cells in series), the relation between current and E.M.F. may be investigated experimentally. Ohm found that they were proportional, or $C \propto E$. This relation is called Ohm's law. In the form of an equation it may be put as

$$C = kE,$$

where k is a constant called the conductivity of the circuit, or

$$C = \frac{E}{R}, \qquad \text{.....(5)}$$

where the constant R is $1/k$, and called the electric resistance of the circuit. This constant is the same whatever be the current or electromotive force, unless the current changes the temperature or physical state of the conductors. It is therefore a characteristic constant of the circuit for a given state and temperature.

If we measure the current through a conductor and the E.M.F. between its ends, we can calculate its resistance from the last equation. Unit resistance thus determined is that of a conductor through which unit E.M.F. drives unit current. On the C.G.S. system, this gives an inconveniently small unit, but the resistance through which one volt drives one ampere, or $10^8/10^{-1} = 10^9$ C.G.S. units, is of convenient magnitude for practical use. It is called the Ohm.

When a standard ohm coil has once been constructed, it

is easy to compare it with other resistances and thus find their value in ohms. The comparison of resistances is best carried out by a piece of apparatus known as Wheatstone's bridge.

Let a long, straight wire be stretched between copper bars along the scale AB (Fig. 42). If the wire is of uniform size throughout, the resistance between two points will be proportional to the length between them. R, a coil of unknown resistance and S a coil of known value in ohms, are inserted in gaps in the copper bars as shown in the figure. A battery of one or more voltaic cells is connected to the bars at the ends of AB, and a galvanometer G is connected to the bar at D between R and S and to a jockey which runs along AB and makes contact with it by a key at any point C.

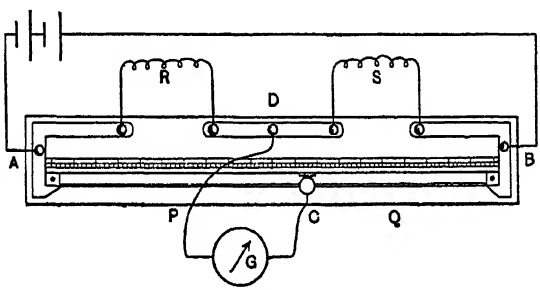


Fig. 42.

The jockey is moved along AB till on making contact with the key at C no current flows through the galvanometer. If E_R and E_S be the E.M.F.'s between the ends of R and S respectively, the current through R and S is the same and equal to

$$\frac{E_R}{R} = \frac{E_S}{S}. \qquad \text{.....(6)}$$

The current through AC of resistance P, is equal to that through CB of resistance Q; hence

$$\frac{E_P}{P} = \frac{E_Q}{Q}. \qquad \text{.....(7)}$$

When C is adjusted so that no current goes through the galvanometer, it means that no difference of potential

exists between D and C, the terminals of the galvanometer wires. Hence $E_R = E_P$ and $E_S = E_Q$. Dividing equation (6) by equation (7) we get

$$\frac{P}{Q} = \frac{R}{S} \text{ or } R = S\frac{P}{Q}. \qquad \text{.....(8)}$$

S is a standard coil of known resistance, and the ratio P/Q is given by the ratio of the lengths AC and CB. Hence R is found. Other forms of bridge are used in which the stretched wire AB is replaced by coils of known resistance. The fundamental principle of them all is the same.

The resistance of a coil is a constant as long as the physical state, including temperature, of the coil is the same. But if the coil be heated, a change is usually observed. In some alloys this change is very small. They are used for resistance coils. But pure metals show an increase of resistance about equal to the change of pressure in a constant volume gas thermometer, that is, for each degree above 0° C. the resistance increases by about 1/273 of its value at 0°. The resistance of a coil of platinum wound on a mica frame, and protected by a tube of glass, porcelain or silica, is used in the platinum resistance thermometer to measure temperature, as already explained on page 58.

Heat and Electricity

If the ends of two wires of different metals are twisted together, and the junction heated above the temperature of the rest of the circuit, an electromotive force is set up, and a current may be obtained. As explained on page 58, this thermo-couple may also be used as a thermometer.

Heating a junction produces a current, and conversely, if a current be passed across the junction of two different metals, a quantity of heat proportional to the total transfer of electricity, that is, to the current multiplied by the time, is evolved if the current pass one way and absorbed if it pass the other. This is called the Peltier effect, from its discoverer.

The reversible Peltier effect, when the rate of heat evolution or absorption is proportional to the current, must be clearly distinguished from the irreversible frictional heat evolved

whenever a current is passed through a resistance. This, known as the Joule effect, is proportional to the square of the current, and, since the square of any quantity, positive or negative, is always positive, heat is evolved in whichever direction the current flows.

Let E be the difference of potential between the ends of a conductor of resistance R, through which a current of strength C is flowing. By the definition of potential difference, E measures the work done by the electric forces when unit quantity of electricity passes. Therefore the total work W is

$$W = ECt,$$

where t is the number of seconds during which the current flows. If no other form of work (such as a chemical decomposition or the driving of a motor) be done, W must, by the principle of the conservation of energy, equal the heat developed expressed in mechanical units (see page 64) or HJ. Hence

$$HJ = ECt.$$

By Ohm's law $E = CR$, therefore

$$HJ = C^2Rt.$$

This relation was used by Joule in his experiments on the mechanical equivalent of heat to determine the value of the constant J. He passed a measured current through a coil of insulated wire of known resistance immersed in the water of a calorimeter. He found J as thus measured to be the same as the value found by the direct conversion of mechanical energy with heat by friction (see page 64).

Ammeters and Voltmeters

The relation between electromotive force, current, and resistance allows a single practical instrument to be used to measure both E.M.F. and current, both volts and amperes.

The coil of a moving coil galvanometer is pivoted, and carries a pointer travelling over a scale. The coil is made of a long length of very fine wire, so that its resistance is high compared with that of anything else in the circuit, and the current through it is consequently proportional to the E.M.F. between its ends. If, then, we connect this coil directly to the terminals of a battery, a very minute current passes through it,

the strength of which is measured by the movement of the pointer and gives the voltage of the battery.

If a strong current is passing through the battery, as for instance when the battery of a motor car is being charged by the dynamo, the strength of that current may also be measured. In series with the dynamo and battery is connected a strip of metal of known but very small resistance, indicated as a heavy spiral line in Fig. 43. Since $E = CR$, for a given resistance the electromotive force between the ends of the strip is proportional to the current flowing. The high resistance galvanometer coil is connected to the two ends of the strip. Its reading is proportional to the E.M.F. and therefore to the current flowing through the strip.

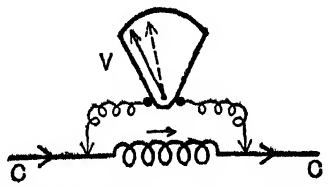


Fig. 43. Volt- and Ammeter.

One end of the galvanometer coil G may be connected with one end of the strip resistance R (see Fig. 44) and with one

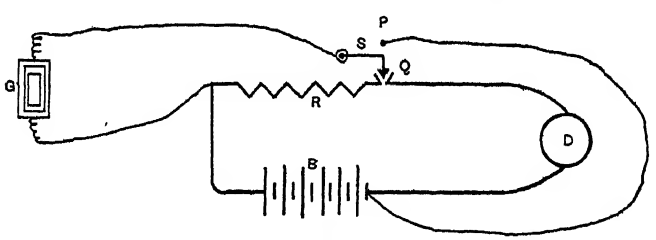


Fig. 44.

pole of the battery B. The other end of the coil is fixed to a switch S, which can be moved so as to make contact with either P or Q. The other end of the resistance R can be joined to one terminal of the dynamo by completing the circuit at Q by inserting the switch S. The other terminal of the dynamo leads to the battery, and completes the main circuit $DBRD$.

When the switch is moved into contact with P, it puts the galvanometer into connexion with both poles of the battery, so that it reads volts. When the switch is pushed into the

gap at Q it completes the main circuit and the galvanometer is connected with the two ends of the resistance R through which the main charging current flows. The reading then gives the amperes flowing through R and therefore round the main circuit.

When the same current is passed through two resistances R_1 and R_2 connected in series, the E.M.F.'s between the ends of R_1 and R_2 will be proportional to their resistances. By connecting a high resistance galvanometer to the ends, first of R_1 and then of R_2, these resistances may be compared with each other. This method is useful with resistances too small to be compared by means of Wheatstone's bridge.

Electromagnetic Induction

We have seen that a current produces a magnetic force, and many vain attempts were made to find the converse effect—the induction of a current by another current or a magnet. Success was first attained by Faraday, whose experiment, described to the Royal Society on 24 November, 1831, opened a new era in electrical science, and proved the starting point for the development of electrical engineering.

"Two hundred and three feet of copper wire in one length was wound round a large block of wood; other two hundred and three feet of similar wire were interposed as a spiral between the turns of the first coil, and metallic contact everywhere prevented by twine. One of these helices was connected with a galvanometer, and the other with a battery of one hundred pairs of plates four inches square, with double coppers, and well charged. When the contact was made, there was a sudden and very slight effect at the galvanometer, and there was also a similar slight effect when the contact with the battery was broken. But while the voltaic current was continuing to pass through the one helix, no galvanometrical appearances nor any effect like induction upon the other helix could be perceived, although the active power of the battery was proved to be great, by its heating the whole of its helix, and by the brilliancy of the discharge when made through charcoal.

"Repetition of the experiment with a battery of one hundred and twenty pairs of plates produced no other effect; but it

was ascertained, both at this and the former time, that the slight deflection of the needle occurring at the moment of completing the connexion, was always in one direction, and that the equally slight deflection produced when the contact was broken, was in the other direction."

It is easy to repeat Faraday's experiment, and to show that a similar result is obtained when a magnet or a coil carrying a current is approaching or receding from another coil. In fact a transient current is induced in a circuit whenever the magnetic force through the circuit is changing. The total number N of lines of magnetic induction through the coil (see p. 94) is called the magnetic flux, and it can be shown experimentally that the total quantity of electricity passing round the coil is equal to the change in the magnetic flux divided by the resistance of the circuit.

The Telephone

One of the most remarkable instances of the use of electromagnetic induction is seen in the telephone. In its simplest form, as invented by Graham Bell and represented in Fig. 45, the telephone consists of a horse-shoe magnet MM with a coil of insulated wire wound over its ends and connected with the land wires at TT. Facing the end of the bar magnet at a short distance away, is a thin iron membrane fixed in a mouthpiece D.

When one speaks into the mouthpiece, the membrane is set in vibration, its mode of vibration depending on the sounds which move it. The approach or recession of the membrane changes the distribution of the magnetic force of the magnet, and therefore the magnetic force through the coil of wire. Transient electric currents are thus set up, the variations in the strength of which depend on the movements of the membrane.

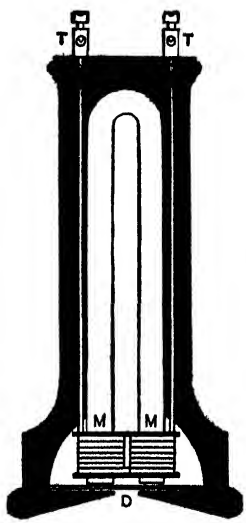


Fig. 45. Bell's telephone.

A similar apparatus is placed at the receiving station. As the transient currents arrive, they change the magnetism of the magnet which they surround. Hence its attraction for the membrane is changed, and the membrane moves accordingly. The surprising result is that the movements of the membrane reproduce accurately those of the membrane at the sending station and thus transmit audible speech.

In more recent forms of instrument, the energy available is increased and the sound reinforced by supplying current from external batteries and by using more complicated apparatus.

Electromagnetic Machinery

We stated on page 108 that the total quantity of induced electricity passing round a coil is equal to the change in the magnetic flux divided by the resistance of the circuit.

This result enables us to give an elementary and approximate theory of the dynamo, the machine which is used to convert mechanical into electrical energy on the large scale.

Let CC' in Fig. 46 denote a coil of wire which is spinning about a diameter as axis with uniform velocity in a magnetic field where the lines of force run in the direction everywhere parallel to AB. When the coil lies along EF, at right angles to the lines of force, the greatest possible number of those lines thread the coil. If H is the strength of the magnetic field or the number of lines per unit cross area, the total flux N is HAn where A is the area of the coil and n the number of turns of wire in it. When the coil has spun through half a revolution, it will again lie in EF, but the flux will now go through it in the opposite direction, and must be written as $-HAn$. The total change of flux is therefore $2HAn$ in

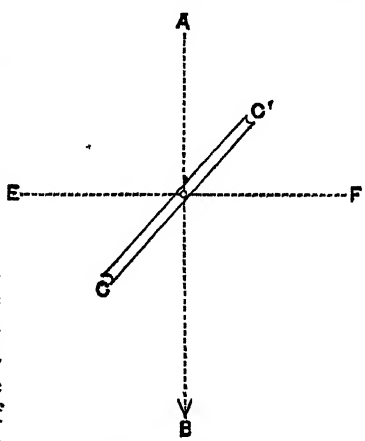


Fig. 46. Earth coil.

each half revolution, and this divided by the resistance, as we have said above, is equal to the total quantity of electricity q which has passed round the coil.

If T be the time of one complete revolution, the time for the half revolution is $\frac{1}{2}T$, and the average strength of current is

$$c = \frac{q}{\frac{1}{2}T} = \frac{2q}{T} = 2qx,$$

where x is the number of revolutions per second or $1/T$. Hence

$$c = 2x \times \frac{2HAn}{R} = \frac{4HAnx}{R}.$$

The current in the coil is reversed each half revolution, and for some purposes these alternating currents may be led off and used as such. But if the ends of the wires of the coil be fixed to copper bars on an ebonite drum on the axis of the coil, the bars will revolve with the coil, and wire brushes pressing against the bars will have their connexions periodically reversed as the current is reversed. Hence the current in the external circuit will always flow in the same direction, and is then called a direct current.

The revolving part or armature of a dynamo consists of an iron core or drum, on which a series of such coils is wound. The current in each coil fluctuates, but the larger the number of coils the more nearly constant is the average current. The field magnets of an alternating current dynamo are excited by a small subsidiary dynamo. The current of a direct current machine may be used not only to do the external work for which the dynamo was designed, but also to excite the electromagnets of the dynamo itself. If the coils of the electromagnets are placed in series with the main circuit, as in Fig. 47, an increase in the external resistance will diminish the current in the magnets and therefore the magnetic field and the electromotive force developed by the dynamo. On the other hand, if the electromagnet coils be wound as a shunt to the external circuit, as in Fig. 48, an increase in resistance in the main circuit will throw more of the current through the magnet coils and increase the E.M.F. For most purposes, as for instance the supply of power and electric light, it is neces-

sary to keep the E.M.F. constant, and then combinations of series and shunt coils are wound on the electromagnets.

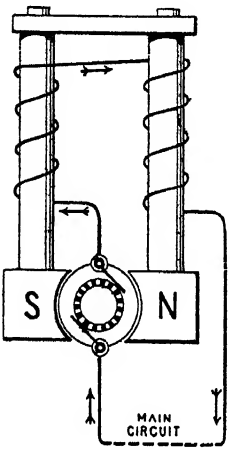


Fig. 47.

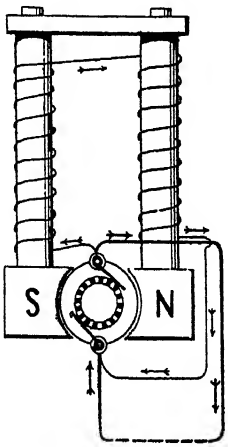


Fig. 48.

The mechanical force acting on a current in a magnetic field is used to obtain mechanical power in electric motors. If a current be passed through the armature of a dynamo in the reverse direction to that of the current the machine itself will yield, it is evident that the armature will tend to rotate. At starting, a large current will pass, but, as the speed increases, a back electromotive force is set up by the usual action of the machine, and the current taken by the motor diminishes. Any dynamo will act as a motor, though the most convenient mechanical arrangements usually differ for the two purposes.

The power conveyed by an electric circuit, that is, the work done per second, is the product of the electromotive force and the current, or Ec. Thus, by increasing E, it is possible to convey a large amount of power by the use of a small current. The conductor required to carry the current with a certain percentage of heat-loss must possess an area of cross-section proportional to the current, and copper is very expensive. Hence it is economical to generate the current

by a high tension dynamo, and convert the power into currents of a safe voltage near the consumer by means of an apparatus called a transformer.

In this instrument, two coils, one of few turns of thick wire, and one of many turns of thin wire, are wound on a core of iron wires or strips, which may either be straight or in the form of a continuous ring. An alternating current of high voltage is passed round the coil of many turns. The magnetism of the core is continually reversed by the alternating current, and the lines of induction through the other coil are reversed continually also. An induced electromotive force, proportional to the number of turns of wire, is obtained, and, as the resistance is low, a large current results. It is obvious that the power of the secondary current cannot exceed that of the primary, and we have as the limit of efficiency

$$E_1 c_1 = E_2 c_2.$$

If the circuit of the secondary coil be not completed, no current flows round it, and no power is taken from the transformer. We should suspect that in these conditions less power is taken from the primary circuit. This is found to be the case.

The induction coil, invented by Ruhmkorff, may be classed as a transformer (Fig. 49). It enables us to obtain an intermittent current of high electromotive force from the direct current given by a few voltaic cells. Two coils are wound on an iron core made of a bundle of soft iron wires to reduce the loss by eddy currents, which would circulate extensively in a solid core. The inner coil is made of a few turns of thick wire, and the outer coil of an immense number of turns of very fine wire wound in sections, most carefully insulated, and so arranged that turns with high differences of potential are separated widely from each other. The primary current, which passes round the inner coil, is interrupted at frequent intervals of time. As a simple means of effecting this, the arrangement used in electric bells may be adopted. The primary current is made to pass from a platinum stud fixed to a spring into a platinum point in contact with it. To the end of the spring is attached a piece of soft iron, which lies near one end of the iron core of the coil. When the current

passes, the core is magnetized, and attracts the piece of soft iron. The spring is pulled away from the point, and the circuit is broken. The spring then flies back, making contact with the point, and the current is established once more.

By some such automatic arrangement, the primary current is made and broken rapidly, and a series of alternately directed electromotive forces are set up in the secondary coil. Since the tubes of magnetic induction produced by the primary current must thread the many turns of the secondary circuit,

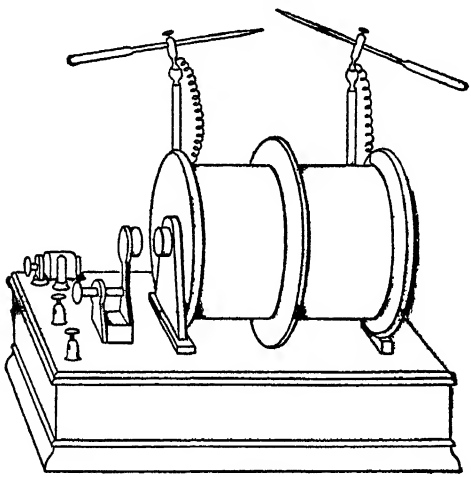


Fig. 49.

the induced electromotive force is very large. Such a coil, then, would furnish an alternating current in the secondary circuit. By the use of a condenser, however, we may increase the electromotive force which arises in the secondary circuit when the primary current is broken. The condenser is made of intervening sheets of tin-foil and paraffined paper, the alternate sheets of foil being connected together. Its opposite systems of plates are connected one with each side of the contact-breaker: in the usual form, one with the platinum point and one with the spring. When the point is in contact with the spring, the condenser is discharged and the core is

magnetized. When the point separates, the condenser is charged through the primary coil, the process taking place by a series of electric oscillations which rapidly demagnetize the iron core. By this means the flux of induction through the coil is destroyed much more sharply than it would be in the absence of a condenser, and the electromotive force on breaking the current is increased greatly.

The use of induction coils has extended considerably of late, owing to their employment in hospitals for the production of Röntgen rays, and their application to many of the more recent branches of physical research.

Electric Units

The forces between electric charges at rest gave us one definition of unit quantity of electricity (see page 91). Its physical dimensions (page 26) may be derived from equation (1)

$$F = \frac{e_1 e_2}{kr^2}.$$

Hence
$$MLT^{-2} = \frac{[e^2]}{kL^2},$$

and
$$[e^2] = ML^3T^{-2}k,$$
$$[e_s] = M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}k^{\frac{1}{2}},$$

where e_s is a quantity of electricity measured in these electrostatic units.

But current electricity gives us another definition of unit quantity, as that quantity passing across any section of a circuit when unit current flows for unit time. To examine this definition we must begin with our fundamental equation for the force between magnetic poles:

$$F = \frac{m_1 m_2}{\mu r^2}.$$

Hence, as above,
$$[m] = M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}\mu^{\frac{1}{2}}.$$

Magnetic force H is the mechanical force F on a pole of unit strength or

$$[H] = \frac{F}{m} = \frac{MLT^{-2}}{M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}\mu^{\frac{1}{2}}} = M^{\frac{1}{2}}L^{-\frac{1}{2}}T^{-1}\mu^{-\frac{1}{2}}.$$

Unit current is defined by the tangent galvanometer (p. 100)

$$H = \frac{2\pi nc}{r},$$

$2\pi n$ is a mere number with no physical dimensions. Hence

$$[c] = HL = M^{\frac{1}{2}}L^{\frac{1}{2}}T^{-1}\mu^{-\frac{1}{2}}$$

and, quantity of electricity (e_m) in electromagnetic units being current multiplied by time,

$$e_m = M^{\frac{1}{2}}L^{\frac{1}{2}}\mu^{-\frac{1}{2}}.$$

Thus we get two values for the physical dimensions of the same thing, namely a quantity of electricity. We can therefore equate e_s and e_m, or

$$M^{\frac{1}{2}}L^{\frac{3}{2}}T^{-1}k^{\frac{1}{2}} = M^{\frac{1}{2}}L^{\frac{1}{2}}\mu^{-\frac{1}{2}}.$$

Therefore,

$$k^{\frac{1}{2}}\mu^{\frac{1}{2}} = L^{-1}T.$$

Now LT^{-1} gives the physical dimensions of a velocity v, which is length divided by time. Hence the dimensions of a velocity are $k^{-\frac{1}{2}}\mu^{-\frac{1}{2}}$ or $\frac{1}{\sqrt{k\mu}}$.

This result must have some physical meaning. Some velocity must be proportional to the reciprocal of the square root of the product of the dielectric constant and the magnetic permeability of an insulating medium. Since no other quantities are involved, it must be some velocity which depends only on the electric and magnetic properties of the medium. If we imagine an electromagnetic wave, whatever that may mean, travelling through a medium with dielectric constant k and magnetic permeability μ, its velocity would depend on these quantities only, and therefore, by the reasoning given above, be proportional to $1/\sqrt{k\mu}$. Clerk Maxwell showed by a mathematical investigation that the velocity of an electromagnetic wave would not only be proportional to this quantity but equal to it, so that we may write

$$v = \frac{1}{\sqrt{k\mu}}.$$

To this fascinating subject we shall return in chapter VII. Here it must suffice to point out that the value of v may be

found by comparing experimentally the electrostatic and the electromagnetic units. v is found to be about 3×10^{10} cms. per second—the velocity of light. Therefore, it seems, light is an electromagnetic wave.

Electrolysis

The result of passing electric currents through solutions was the earliest of their effects to be examined. The subject was again taken up by Faraday, whose experiments showed (1) that the amount of chemical decomposition is proportional to the quantity of electricity which passes, and (2) that the mass of substance liberated is proportional to the chemical equivalent weight of the substance.

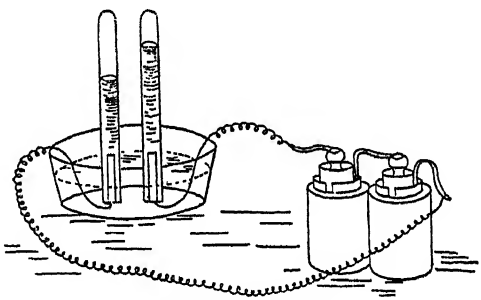


Fig. 50. Electrolysis.

These results may easily be verified. The current from a battery of cells passes through dilute sulphuric acid in an apparatus such as that shown in Fig. 50. By cutting the wire shown in the figure and inserting two more vessels, the same current may be passed also through the solution of a silver salt and through the solution of a copper salt, all arranged in series with the battery and a tangent galvanometer. When the circuit is completed, bubbles of hydrogen and oxygen appear at the platinum plates in the acid, and silver and copper are deposited from their respective solutions. The volume of hydrogen is soon seen to be double that of the oxygen, and these volumes and the weights of silver and copper deposited are found to be proportional to the current as indicated by the galvanometer and to the time it flows.

The weights of the gases can be calculated from the volumes, and it is then seen that the masses liberated are in the ratios H = 1, O = 8, Ag = 108 and Cu = 31·8, that is, in the ratios of the chemical equivalent weights, the atomic weights divided by the valencies. The mass of a substance liberated by unit quantity of electricity, i.e. by a current of one ampere in one second, is called its electrochemical equivalent.

The chemical changes occur only at the contact of the liquids with the metal plates; the bulk of the liquid is unaltered. This is explained by imagining that the substances in solution consist of parts oppositely electrified, and that under the influence of the electric forces these parts, or ions, as Faraday named them, move in opposite directions. Inside the liquid, the opposite ions simply drift past each other, and the number of them in any volume is unchanged. But, at the metal plates or electrodes, the advancing streams find no more opposite ions and are set free. For example, in a solution of silver nitrate the moving ions are positively electrified silver and negatively electrified nitric tri-oxide, which may be written as $Ag+$ and NO_3-. Silver is liberated at the cathode, where the current leaves the solution, and NO_3- at the anode where it enters it. If the anode be a silver plate, the metal is dissolved as $AgNO_3$ and the amount of silver in solution is unchanged. But if the anode be of platinum or carbon, the NO_3- reacts with the solvent water

$$2H_2O + 4NO_3 = 4HNO_3 + O_2$$

and oxygen is liberated.

A similar action explains the decomposition of water when a current is passed through dilute sulphuric acid. The ions are $2H++$ and SO_4--. Hydrogen is liberated as H_2 at the cathode, and at the anode the SO_4+ reacts with water

$$2H_2O + 2SO_4 = 2H_2SO_4 + O_2$$

and again oxygen appears.

The passage of an electric current through a decomposable liquid, or electrolyte, was named by Faraday electrolysis. It consists in the carrying of opposite electric charges in opposite directions by the ions, as water may be passed from hand to hand in a stream of buckets from a lake to a burning house.

Faraday's experiments showed that the mass of an ion deposited was proportional to its chemical equivalent weight. Hence one chemical equivalent of any ion carries the same electric charge, and we see that the charge on a univalent atom is a true natural unit of electricity, just as an atom is a true natural unit of matter. Electricity, like matter, is atomic in its nature.

The polarization at the electrodes prevents the use of an ordinary Wheatstone's bridge to measure the resistance of electrolytes. But it is possible to use an alternating current, with a telephone instead of a galvanometer as indicator. The reciprocal of the resistance between the opposite faces of a unit cube is called the specific conductivity of the solution, and from this quantity the sum of the opposite velocities of the ions may be calculated. The unequal dilution of the solution round the two electrodes gives the ratio of the two ionic velocities, and thus the velocities of the ions themselves may be determined. The results may be verified by watching the motion of coloured ions through a colourless liquid. The velocities are small. With a potential difference of one volt per centimetre, hydrogen travels 0·00323, and silver 0·00058, centimetre per second.

Evidence will be given in the next chapter to show that in these electrolytes the opposite ions possess migratory independence of each other—that a solution of silver nitrate, for example, contains most of its contents as dissociated ions of silver $Ag+$ and nitrate NO_3-.

Conduction of Electricity through Gases

In usual circumstances gases are almost perfect insulators. An electric charge on an insulated metal ball will slowly leak away, but most of this leakage occurs along the surface of the insulating support, whether it be an ebonite pillar or a silk thread. Nevertheless, a true, though very slight, conduction through the surrounding air does take place, and is much increased by various agencies described below. It may be demonstrated by means of a special form of gold-leaf electroscope invented by Mr C. T. R. Wilson, and shown in Figure 84 on page 92. The top of the gold-leaf is fixed to a

metal plate hung from an insulating sulphur bead. The upper end of the bead is fused to a wire, which is put permanently in contact with the terminal of a battery and thus kept at a high potential. The gold-leaf system is charged by momentary contact with this same terminal and then insulated. Its charge is found to leak slowly away, although any conduction over the sulphur bead, since it is kept at a high potential, would tend to increase the charge. The leakage must therefore take place through the air.

When the difference of potential between two conductors in air is increased, a spark will eventually pass. If the electrodes be placed in a glass tube connected with an air pump and the pressure of the air be decreased, the spark broadens out till a fluctuating glow fills the tube. A narrow dark space surrounds the cathode, and, as the pressure is further reduced, the resistance rises and the dark space extends till it fills the tube. The walls of the tube now show a green phosphorescent light, and, by interposing obstacles, this can be shown to be due to the impact of straight rays proceeding from the cathode. They are called cathode rays.

Where the cathode rays strike solid objects such as the glass walls of the tube, or an anticathode of aluminium placed within the bulb (see Fig. 51), another kind of ray, discovered by Röntgen, is generated. These are the famous X-rays. They affect photographic plates and illumine fluorescent screens. They are absorbed by matter in rough proportion to its density, and thus coins in a purse or bones in a hand stop more of the rays than do the envelopes, and may be made visible by their shadows.

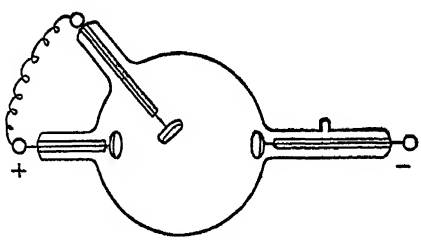


Fig. 51.

The cathode rays carry negative electric charges and are deflected from their straight paths by electric and by magnetic fields. Sir J. J. Thomson (now Master of Trinity College, Cambridge) investigated this effect by means of an apparatus

resembling that shown in Fig. 52. A thin pencil of rays from the cathode C is passed between two plates D and E connected with the opposite poles of a battery. The electric force deflects the cathode rays on the fluorescent screen and shifts them from p to p'.

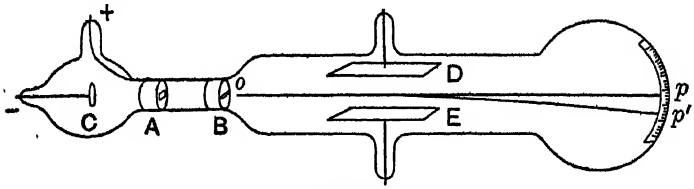


Fig. 52.

The effect is clearly the same as that of gravity on a body thrown horizontally. If f be the electric force and e the charge on a particle of mass m, the mechanical force is fe, and the acceleration fe/m, and, if the displacement at DE be S_e, we have, by the usual laws of falling bodies (p. 31),

$$S_e = \tfrac{1}{2}at^2 = \frac{1}{2}\frac{fe}{m}t^2.$$

During the time t the particle is moving horizontally with uniform velocity v and traverses a length l equal to vt. Hence t^2 is l^2/v^2 and

$$S_e = \frac{1}{2}\frac{fe}{m}\frac{l^2}{v^2}.$$

Next, instead of the electric field, a magnetic force H is applied by placing the whole tube between the poles of an electromagnet. The moving particle with its charge is equivalent to a current, and thus the magnetic deflection S_m may be calculated. A mathematical investigation shows that it is given by

$$S_m = \frac{l^2He}{2vm}.$$

Thus, by measuring the electric and the magnetic displacements, we get two relations between e/m and v, and can determine both these quantities. Thomson found that, while v varied from 2×10^9 to 4×10^9 cms. per sec., the quantity e/m was constant whatever was the pressure or nature of the gas and whatever metal was used as electrode. The most

recent values for e/m are about $1{\cdot}77 \times 10^7$ electromagnetic units. In liquid electrolytes the passage of one electromagnetic unit of electricity evolves 10^{-4} gram of hydrogen. Here m/e is 10^{-4} and e/m is 10^4, and it follows that, for the particles of cathode rays, e/m is about 1800 times greater than in liquids and m/e about the 1/1800 part of its value in liquids. Thomson gave reasons to believe that e was the same as the charge on single ions in liquid electrolytes, and this conclusion was soon verified in other ways. Therefore the particles or corpuscles in cathode rays possess a mass only 1/1800 of that of a hydrogen atom.

Here, then, in Thomson's corpuscles we have particles much smaller than the least of the chemical atoms. They are the same whatever the nature of the substance from which they are derived. The age-long dream of a common basis for matter seems coming true. Moreover, each particle is associated with a unit negative charge of electricity, and may perhaps be only another way of regarding that unit of electricity. Incandescent bodies emit light, and, as we shall see in chapter VII, light consists of electric waves. Electric waves must arise from moving electric charges, and thus the ultimate basis of matter must be electrical. Matter and electricity may ultimately prove but different manifestations of a same underlying something. Thomson's corpuscle, identified with a negative electric unit as a basis of matter, is now called an electron.

Positive Rays

The cathode rays carry a negative charge and come from the negative electrode. The positive electrode or anode also emits rays which were discovered by Goldstein and have been investigated by Thomson and by Aston. The mass of the cathode ray particles is independent of the nature of the residual gas in the tube, but this is not so with positive rays. They contain particles with masses equal to those of different chemical atoms.

Aston passes positive rays through two slits and then, by electric and magnetic deflection of the narrow beam so obtained, forms a spectrum in which rays with a constant

value of m/e fall in the same place. The mass spectrum thus obtained is photographed, and is found to consist of lines characteristic of different elements. The absolute values of the atomic mass are found by comparison with the lines of known elements such as hydrogen and oxygen which occur as impurities. One such mass spectrum is shown on Plate II.

With this apparatus Aston has shown that many elements with atomic weights not whole numbers consist of mixtures of two or more kinds of atom. Thus chlorine (35·46) gave a spectrum of two lines, corresponding to atomic weights of 35 and 37. Except hydrogen (1·008) all the elements examined consist of atoms with weights exactly whole numbers on the scale which takes oxygen as 16.

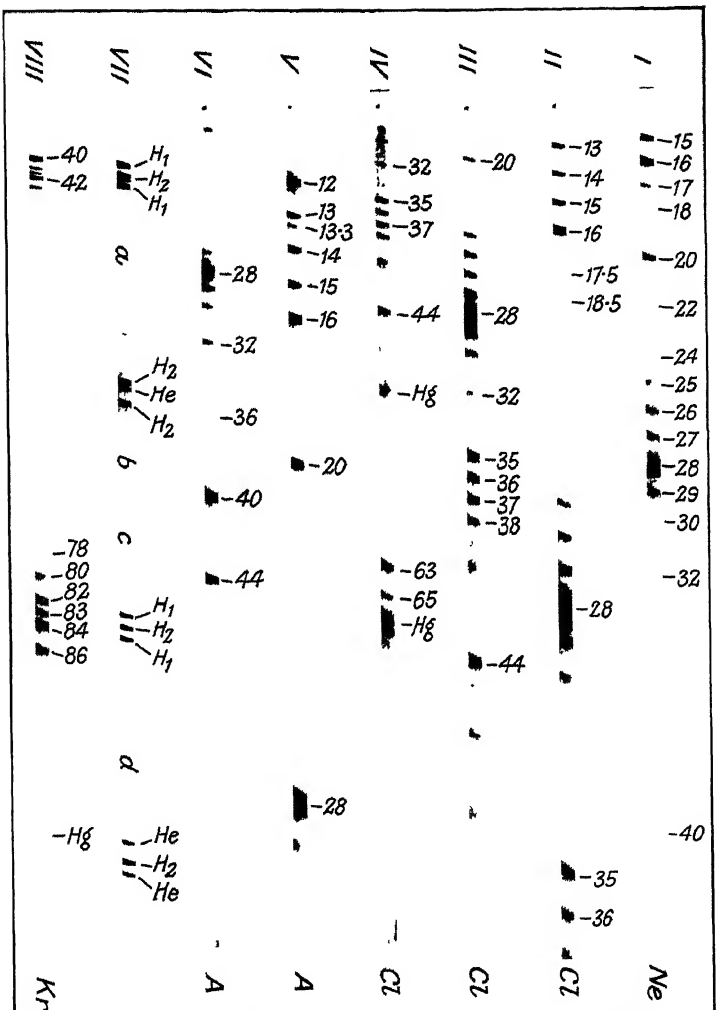
These different atoms of the same chemical element were called isotopes by Soddy. In some cases they have now been separated. Thus Soddy has prepared by radioactive means lead differing in atomic weight from lead as usually obtained.

Electronic Theory of Conduction

The discovery of the cathode ray particle, and its identification with the concept of a negative electron, led at once to a new view of the process of metallic conduction.

In liquid electrolytes, a current can only pass by the aid of moving ions at least as large as chemical atoms, and more probably complex structures composed of an atomic nucleus with a surrounding envelope of solvent. But in metals no chemical changes accompany an electric current; the only obvious effect is the development of heat proportional to the resistance. Here there is no movement of molecules or atoms, and we must imagine that the negative corpuscles or electrons slip from atom to atom. The current is carried by negative electrons only, and the heat developed is the equivalent of the friction involved in their movement.

Faraday and Maxwell showed the importance of the dielectric or insulating medium. When an electromotive force is set up, an electric impulse spreads through the medium, as we shall see in chapter VII, with the velocity of light. A dielectric medium is elastic, it tends to oppose the force which acts on it, as a spring opposes a force which extends or



TYPICAL MASS-SPECTRA

compresses it. But where a conducting channel is open, as along a metal wire, the mechanism slips, energy is degraded into heat, and what we call a current flows, its direction being guided along the conductor where alone this slipping can occur. Thus the energy travels through the surrounding dielectric medium; it is only the dissipation of energy which constitutes an electric current.

A flame or the hot gases therefrom will discharge an electrified body, showing that they contain ions. Near an incandescent solid also the air is ionized. As a platinum wire is heated, it begins to emit positive ions of atomic size or larger at a low red heat. At higher temperatures negative particles come off, ultimately in large excess. In a vacuum the negative leak from filaments of platinum, carbon or tungsten is very large, the ions being the sub-atomic electrons. A positive charge is, of course, left behind, and its attraction will finally prevent further emission. A state of equilibrium is thus reached, analogous to that between a liquid and its vapour. Raising the temperature will cause more electrons to come away, and increase their numbers in the surrounding space just as it increases the vapour pressure of a liquid.

This effect must be important in cosmical processes. The photosphere of the sun contains masses of glowing carbon, which will shoot forth electrons till the attraction of the positive charge left behind prevents further emission. A local elevation of temperature will cause a new flood of electrons to dart through space. Perhaps the Aurora Borealis may be due to these corpuscles streaming through the upper region of the earth's atmosphere.

The emission of electrons from incandescent filaments has become of great industrial importance in recent years. It forms the basis of the thermionic valves now so extensively used in wireless telegraphy and telephony. This application will be dealt with in chapter VII.

CHAPTER V

CHEMICAL ACTION AND EQUILIBRIUM

Physical and Chemical Equilibrium

In the preceding chapters we have described the chief facts and explained the accepted principles in chemistry, dynamics, heat and electricity. We are ready now to apply our knowledge to the general problem of equilibrium and other questions which lie in what is usually regarded as the borderland of physics and chemistry. Nature knows no division of subjects: the separation of studies is a mere matter of human convenience. Indeed that separation, though inevitable, often delays an advancement in learning for which the methods and outlook of two different sciences are needed. Hence, when the results of one science are carried over and applied to the facts of another, there often follows a great revelation of new knowledge.

A dynamical system comes to rest when its potential energy is a minimum—the pendulum when at its lowest point, a falling stone when it reaches the ground. When thermal energy is involved, we have seen (p. 77) that the available energy tends to diminish so that equilibrium is reached when it is at a minimum.

Available energy and similar quantities, then, are analogous to potential energy in dynamics. Hence they are sometimes called thermodynamic potentials. They depend on the whole nature and state of the system; on its chemical constitution, on the external variables of temperature and pressure, and on surface tension and electric charge if these influences be appreciable.

Now, if we have but one unknown quantity x, a single simple equation will give us its value. If we have two, x and y, the elementary algebra of simultaneous equations has taught us that two independent equations are needed. If we can construct but one equation involving x and y, we cannot solve it; but, even with only one equation, if we choose a

value for one of these unknown variables, the value of the other is fixed also. We can freely choose a value for either x or y, but not for both; the system is said to have one degree of freedom.

And so in the general case. If we can find as many independent relations between them as we have independent variables, the system is fixed and determinate. If we have fewer equations, it possesses "freedom," and the number of "degrees of freedom" is measured by the excess of the number of variables over the number of equations.

The system will be in equilibrium when the appropriate potential is a minimum. For instance, if water be in contact with its saturated vapour in a closed vessel, water will evaporate or steam condense till the thermodynamic potential, ψ, at constant volume of the whole system is a minimum.

Let the curve of Fig. 53 represent the relation between the mass m of water substance in the state of vapour and the potential ψ. It is clear that at P, where ψ is a minimum, a small change in m does not affect the value of ψ, since at P the ψ-curve is horizontal. Hence equilibrium is reached when the rate of change of ψ with m, which may be written as $d\psi/m$, vanishes.

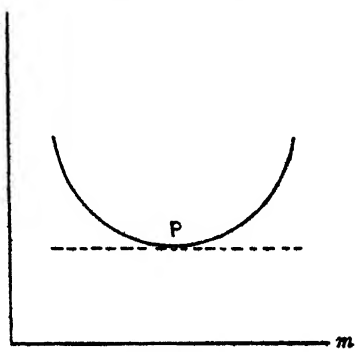


Fig. 53.

Now let us consider both liquid and vapour. Let ψ_1 refer to liquid and ψ_2 to vapour. If a small mass dm evaporate, ψ_1 decreases by a small quantity $d\psi_1$, and ψ_2 increases by $d\psi_2$, and for equilibrium we have

$$\frac{d\psi_2}{dm} - \frac{d\psi_1}{dm} = 0 \quad \text{or} \quad \frac{d\psi_2}{dm} = \frac{d\psi_1}{dm},$$

which may be written for short $\mu_2 = \mu_1$. Hence there is equilibrium between two phases (e.g. liquid and vapour) of one component (e.g. water substance) when $d\psi/dm$ is the

same for both. The quantity $d\psi/dm$, the rate of change of ψ per unit transfer of mass from one phase to the other, is called the chemical potential, and two phases are in equilibrium when their chemical potentials are equal.

The Phase Rule.

Water substance, the one component of the system we have considered, can exist in three states or phases, solid ice, liquid water and vaporous steam. Any two of these phases are in equilibrium when their chemical potentials are equal, just as two communicating vessels of air are in equilibrium if their pressures are equal. Hence we can form one equation between the chemical potentials of two phases, and the chemical potentials involve the chemical constitution and the two external variables temperature and pressure. If all three phases, ice, water and steam, are in equilibrium together, we have three equations $\mu_1 = \mu_2$; $\mu_2 = \mu_3$ and $\mu_3 = \mu_1$, but the third follows from the first two, and is not an independent relation: for three phases we have two equations. And so on: for r phases we can form $r - 1$ independent relations.

If, instead of one component, water-substance, we have two independent components, such as water and common salt, the result will apply to each. And, if we have n components, the total number of independent equations is $n(r - 1)$.

Now let us consider the number of unknown variables. If we have one gram of a solution of salt in water, and know that one-tenth of a gram is salt, we can say at once that nine-tenths of a gram is water; the amount of water is known as soon as the mass of salt is fixed. Hence for two components we have only one independent variable, and for n components there would be $n - 1$ variables. This holds good for the components in each phase, and so for r phases the number of variables fixing the composition is $r(n - 1)$. Adding the two general external variables temperature and pressure, and ignoring possible variables dependent on other quantities such as electric charge, we get the total number of variables

$$r(n - 1) + 2.$$

Now, as we have seen, the number of degrees of freedom F is equal to the excess of the number of variables over the number of independent relations between them, or

$$F = r(n-1) + 2 - n(r-1),$$

or

$$F = n - r + 2. \qquad \ldots\ldots(1)$$

This is Willard Gibbs' Phase Rule.

One-Component Systems

If one component, such as water substance, alone be present, n is unity and the equation gives

$$F = 3 - r, \qquad \ldots\ldots(2)$$

or the number of degrees of freedom is three, less the number of phases in equilibrium. If we collect three phases, that is, if ice, water and vapour be in equilibrium together, r is 3, F is zero, and the system has no degrees of freedom but is fixed and determinate. That means that the three phases can coexist only at one value of the temperature and one value of the pressure, namely at the freezing point of water under its own vapour pressure. The freezing point of water under the pressure of one atmosphere is called 0° C., and the freezing point is lowered by about 0°·007 for each additional atmosphere (see p. 129). The vapour pressure at the freezing point is 4·7 mm. of mercury—a small fraction of one atmosphere. Hence the temperature of equilibrium is about + 0°·007 C., and at that temperature and at a pressure of 4·7 mm. and there alone, ice, water and vapour can coexist in equilibrium. This is called the triple point.

If this triple mixture be heated, ice melts. As long as any ice remains, there is and can be no change in temperature or pressure, but, when the last morsel of ice has gone and only water and vapour are left, we have two phases. Equation (2) then shows one degree of freedom, that is, equilibrium can exist at any temperature within limits, provided the pressure is at the corresponding saturation value.

A phase rule diagram, drawn between temperature and pressure, will make these relations clearer (Fig. 54). Ice, water and vapour coexist at the triple point O, which has definite

values of t and p, so that the system is nonvariant. When the ice has melted, water and its vapour can coexist with one degree of freedom in a monovariant system. As a nonvariant system must be represented by a point, a monovariant system must be shown by a line on which we can choose either t or p arbitrarily, but then p or t is fixed. For instance, if we adjust t to 50° C., equilibrium can only exist at a pressure of 92·3 mm. of mercury. The equilibrium represented by OA reaches from the triple point to the critical point, beyond which no liquid can exist.

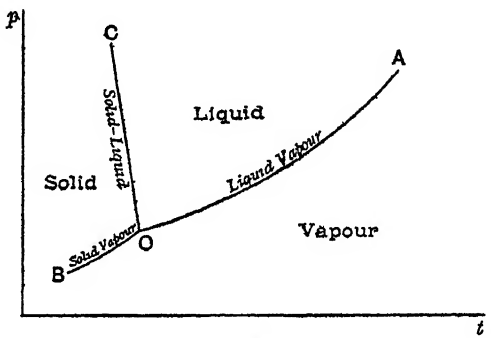


Fig. 54.

Beginning again at O, let heat be abstracted. Water freezes, and finally we are left with ice and vapour. We pass along the monovariant line OB giving the vapour pressures of ice below its melting point.

Again, if the triple mixture at O be compressed, vapour is condensed, and in the end we are left with ice and water only, to trace the monovariant equilibrium represented by the curve OC. A very large increase of pressure is needed to lower the melting point by 1° C.; hence the curve is very steep.

The slope of all these curves is determined by the latent heat and the alteration of volume of the change from one phase in equilibrium to the other. The relation is (p. 81)

$$\frac{dp}{dt} = \frac{L}{T(v_2 - v_1)}$$

so that (p. 82), as in the case of water and its vapour, if we know L the latent heat, T the temperature of transition, and $v_2 - v_1$ the change in volume, we can calculate dp/dt, that is, the slope of the tp curve. For instance, one gram of ice contracts by 0·084 c.c. on melting; thus $v_2 - v_1 = -0{\cdot}084$. Near the freezing point 273° A., i.e. 273° on the absolute scale, the latent heat of fusion is 80 thermal units or $80 \times 4{\cdot}2 \times 10^7$ mechanical units. Hence the increase of pressure needed to change the temperature of equilibrium by 1° C. is

$$dp = -\frac{80 \times 4{\cdot}2 \times 10^7}{273 \times 0{\cdot}084} = -1{\cdot}464 \times 10^8 \text{ dynes per sq. cm.}$$
$$= -144{\cdot}5 \text{ atmospheres.}$$

The negative sign means that as p increases, t decreases, that is, that the melting point is lowered by pressure, and ice at 0° C. is melted if the pressure be raised.

This can be well shown by hanging two heavy weights over a lump of ice by a fine copper wire. The ice is melted by the pressure of the wire, which cuts its way through the lump, leaving solid ice behind it where the pressure is relieved.

Two-Component Systems

With salt and water we have a system containing two independent components. In the phase rule equation n is 2 and $F = 4 - r$, so that to get a nonvariant system we must collect together four phases.

This is quite possible. There are two solids available, crystals of ice and crystals of salt, one liquid, the saturated solution of salt at the freezing point, and one vapour—that of water.

In the phase rule diagram, we must introduce, as in Fig. 55, a third axis c to represent the composition of the liquid phase, the concentration of the solution. Hence the diagram must now represent a three-dimensional solid model, with three axes at right angles to each other—t, p and c. Oc represents a line at right angles to the plane of the paper.

The point O represents the nonvariant system—the equilibrium of the four phases. If heat be added, ice melts and salt dissolves in the water so formed, so that the concentration of the solution, the temperature and the pressure remain

constant. This constancy of melting point used to be thought the sure and certain sign of a pure chemical compound. Hence these mixtures were called cryohydrates by their discoverer Guthrie, from the idea that they were compounds of salt and ice. We now see that their constancy of melting point indicates not a compound but a mixture containing two distinct solid phases in equilibrium with each other, with the saturated solution and with the vapour. These mixtures form very useful thermostats, that is, baths of

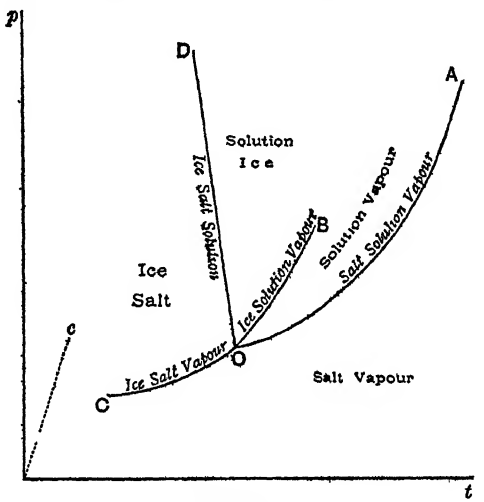


Fig. 55.

constant temperature. A very soluble salt will give a very low cryohydric point, e.g. calcium chloride is in nonvariant equilibrium with water at $-55°$ C. while sodium chloride gives the usual freezing mixture of ice and salt at $-22°$ C., near zero on the Fahrenheit scale.

If heat be passed into a cryohydric mixture, either ice or salt will fail first. If ice, then we get left with the three phases, salt, solution, vapour, and pass along the monovariant curve *OA*, giving the vapour pressures of solutions saturated at different temperatures. If salt fail first, and we

are left with an excess of ice, the further melting of it will dilute the solution, and we pass along the second monovariant curve *OB*, which gives the freezing points of solutions of continually increasing dilution till, on melting an indefinite excess of ice, pure water is reached at *B*. Here the diagram, if it were drawn in full, would reproduce Fig. 54.

If heat be abstracted from the cryohydrate, all liquid will vanish, and we traverse *OC*, the vapour pressure curve of mixtures of ice and salt. Finally, if the pressure at *O* be increased, we pass along *OD*, where there is no vapour and the equilibrium is that between ice, salt and solution.

We now see that many experimental investigations, such as the measurement of vapour pressure, the determination of the freezing point of solutions or the solubility of a salt, are in their essence, the tracing of a curve in a phase rule diagram. They are all cases of equilibrium, the conditions of which are completely described theoretically by the phase rule and the latent heat equation.

The complete case of mutual solubility of two bodies is well illustrated by certain alloys, of which we will take that of silver and copper as an example. Pure copper melts at 1081°·5 and the addition of silver lowers the melting point in a regular and monovariant manner as shown by the curve on the right in Fig. 56. On the other hand, silver melts at 960°, and on adding copper the melting point again is lowered as represented by the curve on the left. The curves cut each other at a point representing a state in which there exist (1) crystals of silver, (2) crystals of copper, (3) the saturated solution of either in the other, and (4) the vapour, the number of phases being that required to give a nonvariant system.

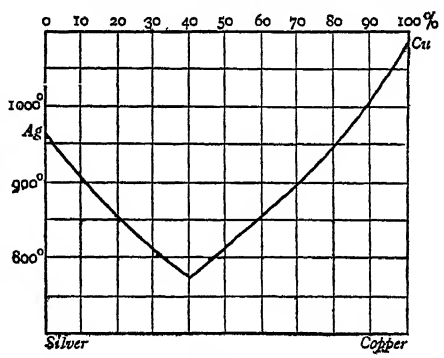


Fig. 56.

The alloy has then the composition 60 atomic per cents. of silver and 40 of copper; it has, of course, a constant melting point, namely 777°; and, owing to its regular texture, is called a eutectic alloy. Clearly it resembles the cryohydric mixtures formed with water and salts. If a fixed alloy of another composition, such as that represented by the line 20 in Fig. 56, be allowed to cool, crystals of pure metal, in this case silver, first appear. The residual liquid thus becomes richer in copper, and this goes on till it reaches the composition of the eutectic alloy, at which the whole of the residual liquid solidifies. In Fig. 57, which represents an alloy of gold and aluminium, the large crystals are pure gold and the uniform eutectic alloy fills evenly the space between them. When one component is nearly pure, that is, is present in large excess, the alloy with other materials only appears as a thin line of cement joining the crystals together. This is shown in Fig. 58, which represents crystals of cadmium connected by thin channels containing impurities. Similar appearances are shown by commercial iron which is an alloy of iron and carbon. The size of the crystals depends also on the rate of cooling, being larger if the cooling is slow. By these considerations the physical structure of simple alloys, as revealed by a microscope, has been explained.

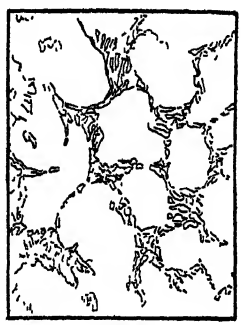


Fig. 57.

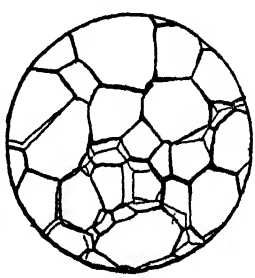


Fig. 58.

If the two components unite to form a chemical compound, the melting point of that pure compound will be lowered by adding excess of either constituent, just as the melting point of either pure component is lowered by the other. Hence the point representing the compound must be a maximum and the curve must fall away on each side. This is shown by the point *b* in Fig. 59, which illustrates the alloys of copper and antimony.

Hitherto we have dealt only with solids of definite composition, but in many cases the composition of a solid phase may vary continuously like that of a solution. For instance, potassium can replace sodium in soda alum by insensible stages, the form of the crystals remaining unaltered. Such solids are called mixed crystals or solid solutions. Their melting and solidifying points are not always identical, and the theory of these changes is complex. Nevertheless, by its help, most of the phenomena, for instance, of the solution of

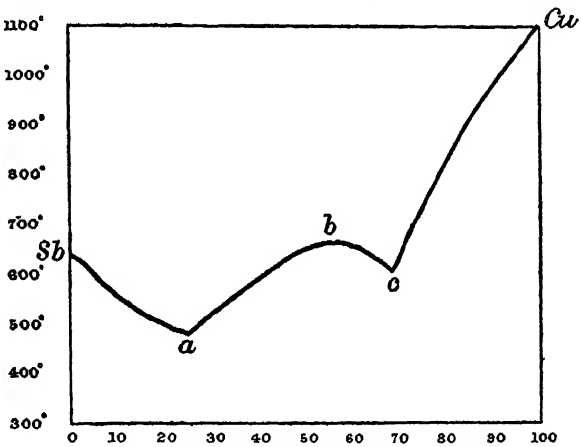


Fig. 59.

carbon in iron have been explained—phenomena which appear in the various and important properties of different varieties of iron and steel.

Solution

We have already used solutions in illustration of the principles of equilibrium, and a knowledge of the obvious properties of solutions in water of sugar or salt or of compressed carbon dioxide gas may be assumed as common to most people. These substances or solutes dissolve in the solvent water to form homogeneous, one-phase systems from which the components may be separated by evaporation. While in solution, they modify the physical and chemical properties of the

solvent. The density, viscosity, surface tension, freezing and boiling points, etc., are all affected, and the chemical reactions of the solution are those characteristic of the dissolved substance, as well as of water.

Very great differences in solubility are known, ranging, indeed, from the apparent insolubility of gold in water to the complete mutual miscibility of water and alcohol.

The dimensions of the dissolved particles also range between wide limits, though Graham divided solutions into two distinct groups, those of crystalloids and colloids. Crystalloids, bodies which usually come out of solution in definite crystals, when in solution, readily pass through ordinary membranes and form single-phase, homogeneous liquids. Colloids, somewhat gelatinous in the solid state, disperse in solvents to form two-phase mixtures, liquids in which are suspended particles either visible in a microscope or able to make their presence known by polarizing light, or scattering its rays when viewed by oblique light in an "ultra-microscope."

As instances of substances which form colloid solutions in water, we may give starch, albumen, gelatine, many dyestuffs, ferric hydroxide, and arsenic sulphide, while, by passing electric sparks between electrodes of certain metals, such as platinum, under water, colloidal solutions of the metals may be obtained.

Many of these solutions are somewhat unstable, and may be precipitated by minute quantities of electrolytes. When finely divided mud is carried down by a river, it is sometimes precipitated by the first contact with sea water. Such results as this probably show the effect of electric ionic charges on surface energy.

Many different views have been held about the nature of solution. It is evident that solubility means a decrease in available energy on solution, but how that degradation of energy is effected is another question. It has been suggested that salt dissolves in water as a liquid evaporates into an empty space, that it is a case of chemical action, or that surface tension is the determining factor. Possibly these statements represent different aspects of the same fundamental process.

We now pass to consider in more detail some of the properties of solutions.

Osmotic Pressure

We have seen (pp. 47, 72) that the gaseous state of matter shows properties and relations that are simpler than those of liquids or solids, because in gases the molecules are scattered throughout the whole space available, and, in general, are so far from each other that the forces between them are negligible (p. 70).

In a dilute solution, the particles of the dissolved substance, or solute as it is called, are also far separated from each other. Hence those properties of dilute solutions which depend on the solute particles and their interaction may be expected to show some of the simplicity of the properties of gases, though other properties of the solutions, which depend on the connexion between solute and solvent, will resemble those of liquids, and show no such simplicity.

This subject was opened up by the discovery of what is called osmotic pressure. It has long been known, as we have seen above, that certain liquids and certain solids classed as crystalloids dissolved in them pass easily through organic membranes, while other solids called colloids will not pass.

These phenomena are of obvious importance in animal and vegetable cells, and it was in a botanical laboratory that the fundamental discovery was made by Pfeffer in 1877, while its physical and chemical importance was pointed out by the Dutch physicist Van't Hoff in 1885.

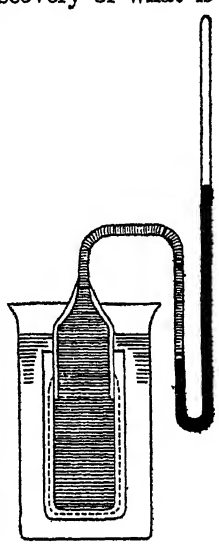


Fig. 60. Osmotic cell.

Fig. 60 shows a porous pot of unglazed earthenware, six to eight centimetres high and two to three in diameter. A glass tube which just fits inside the pot is fixed watertightly by sealing wax. The glass tube is drawn off into a pressure gauge as shown. The pot is filled with the solution of a substance such as potassium ferrocyanide, and dipped into another solution, such as one of copper sulphate, which forms

an insoluble precipitate when it touches the first solution. As the liquids slowly creep into the walls of the pot, they come in contact and form an insoluble membrane of copper ferrocyanide which, in a certain (or rather, uncertain) number of lucky attempts, completely encloses the pot.

The solutions are then washed out, and the apparatus is complete. The membrane of copper ferrocyanide is found to be permeable to water, but not to solutions. It is therefore called a semipermeable membrane. With a good membrane even crystalline salts may be retained, and substances such as cane sugar are completely held up. Hence, if the pot be surrounded by pure water, and filled with a solution of cane sugar, water will enter, and Pfeffer found that this process went on till an equilibrium pressure was reached which was proportional to the concentration of the sugar in dilute solution.

Now the concentration of a dilute solution is the reciprocal of the volume v in which unit mass is dissolved. Hence, Pfeffer's result may be stated by saying that, in a dilute solution, the osmotic pressure is inversely proportional to the volume. But that is Boyle's Law of Gases applied to dilute solutions.

Again, if solutions of two different substances be prepared, so that the numbers of gram-molecules (see p. 67) are the same in unit volume of each, that is, if they are of equimolecular concentration, it will be found that at the same temperature they possess equal osmotic pressures. This is Avogadro's Law. The osmotic pressure also increases with temperature at the same rate as does gaseous pressure.

Once more, it is found that an osmotic pressure equal to that of one atmosphere is given by a solution which contains one gram-molecule in a volume of about twenty-two litres. This shows that the absolute value of the osmotic pressure is the same as the pressure of a gas which contains the same number of gram-molecules in the same volume (see p. 45).

It is now clear that there is a very close analogy between gases and dilute solutions. It does not follow that the physical cause of the pressures is the same in the two cases, indeed thermodynamic theory shows that the analogy is merely a consequence of the particles in each case being beyond each

other's sphere of force, and will hold good whatever be the nature of the interaction between solvent and solute.

The theoretical importance of osmotic pressure, indeed, lies chiefly in this fact, that it enables thermodynamic reasoning to be applied to solutions. If we imagine, as in Fig. 61, the bottom of an engine cylinder to be made of a porous wall containing a semi-permeable membrane, with a solution within and pure solvent in contact without, it can be imagined to do work as Carnot's imaginary engine does.

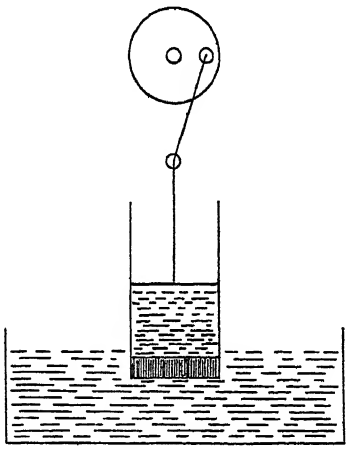


Fig. 61.

Besides natural membranes and the artificial ones we have described, there are semipermeable membranes of the simplest and most perfect kind in the free surface of a volatile liquid such as water in which some non-volatile substance like common salt is dissolved. Again, the crystals of ice which are formed by freezing a dilute salt solution, consist of pure water-substance and contain no salt. Hence the surface of a crystal of ice is also a perfect semipermeable membrane, and thermodynamic results may be applied to evaporation and also to freezing.

Let us imagine a tall cylinder to contain solution and to possess a semipermeable bottom. Solvent will enter and the level of the solution rise till the pressure on the bottom within is equal to the osmotic pressure P. There is then equilibrium. Now $P = h\rho g$, where h is the height of the column, ρ the density of the liquid, and g the acceleration of gravity (see p. 31). The pressure p of the vapour at the level of the solvent outside is greater than its pressure p' at a height h by

$$p - p' = h\sigma g,$$

where σ is the density of the vapour assumed uniform over

a small height. The pressure of the vapour at the surface of the solvent is equal to the vapour pressure of the solvent, and at a height h, equal to that of the solution, for if not there could be no equilibrium, a circulation would be set up, and we could get work from the internal reactions of an isothermal system, contrary to the principles of thermodynamics (see p. 75). Hence

$$p - p' = \frac{P}{\rho g}\sigma g = P\frac{\sigma}{\rho},$$

or

$$P = (p - p')\frac{\rho}{\sigma},$$

which gives a relation between the osmotic pressure and the lowering of vapour pressure of a solvent by dissolving in it a solute.

By imagining a cycle of operations in which a gram of ice from a solution is frozen, removed from the solution, melted and returned to the solution, we may apply the latent heat equation, and get a relation for the depression t of the freezing point T in the form

$$t = \frac{PT}{L\rho},$$

where L is the latent heat of fusion and ρ, the density of the liquid, is the reciprocal of the volume of unit mass (see p. 81).

In water, $L = 80 \times 4{\cdot}18 \times 10^7$, $\rho = 1$ and $T = 273°$ A. If we have a solution containing one gram-molecule per litre, $P = 22{\cdot}2$ atmospheres or $2{\cdot}25 \times 10^7$ C.G.S. units, and we get

$$t = \frac{2{\cdot}25 \times 10^7 \times 273}{80 \times 4{\cdot}18 \times 10^7} = 1°{\cdot}85 \text{ C.}$$

The depression of freezing point in other solvents for the same concentration of solution may be calculated in the same way. The following is a comparison of these calculated results with Raoult's observed numbers:

	T	L	t calculated	t observed
Water	273	80	1·85	1·85
Acetic acid	290	43·2	3·88	3·86
Formic acid	281·5	55·6	2·84	2·77
Benzene	277·9	29·1	5·30	5·00
Nitrobenzene	278·3	22·3	6·95	7·07

Thus the experiments may be regarded as a striking confirmation of the theory.

The osmotic pressure of a solution is a very difficult quantity to measure accurately by direct experiment, but the relations given above show that it may be determined indirectly from either the lowering of vapour pressure or the depression of the freezing point of a solution compared with its solvent.

One application of these results is of great importance in chemistry: they enable us to estimate the molecular weight of a substance in solution. A solution which contains one gram-molecule per litre may be taken as giving a depression of the freezing point equal to that shown in the last table, and, for dilute solutions, the depression is proportional to the molecular concentration. Hence, by making up a dilute solution with a known number of grams per litre and measuring the freezing point (or the lowering of vapour pressure) we can tell at once the number of gram-molecules per litre, and thus calculate the molecular weight. As we saw on p. 45, vapour density determinations give the molecular weight as a gas, and these new methods give the molecular weight in solution.

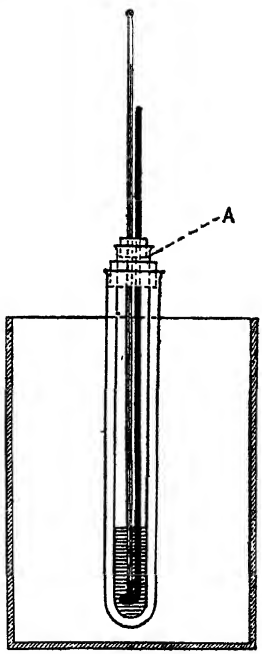


Fig. 62.

For instance, suppose we wish to find the molecular weight of alcohol when dissolved in water. A solution is made up containing, say, 2·2 grams in 100 c.c. A convenient quantity is placed in a glass tube (*A* of Fig. 62). This tube is fixed inside another, rather larger, tube and the whole immersed in a freezing mixture of ice and salt. The solution is kept stirred and the thermometer watched. Generally it sinks below the freezing point, till suddenly some ice is formed. The temperature then rises, and the highest point reached is recorded. Since, owing to the undercooling, a considerable quantity of ice is removed from solution on

freezing, the concentration of the solution which remains is greater than that prepared, and a correction must be applied. Suppose we find the corrected concentration is 2·3 grams in 100 c.c., or 23 grams per litre, and the depression of the freezing point 0°·925.

The molecular concentration is $\frac{0{\cdot}925}{1{\cdot}85} = 0{\cdot}50 = \frac{1}{2}$, and the molecular weight consequently twice 23, or 46. This shows that the formula weight, $C_2H_5.OH = 46$, is the molecular weight in solution.

Osmotic Pressure of Electrolytes

As long as we examine solutions which do not conduct electricity, the measurement of the osmotic pressure, either directly, or indirectly by freezing point or vapour pressure, shows that the molecule of the solute is that represented by the chemical formula or some multiple of it. But solutions of electrolytes show different results.

Any one of the osmotic effects, either the osmotic pressure, or the lowering of freezing point or of vapour pressure, for a dilute solution of a binary electrolyte such as potassium chloride, is twice as great as the concentration of the solution would suggest. Since these effects are proportional to the number of pressure-producing particles, this doubled value indicates that each molecule of KCl has dissociated into two parts, each producing its own osmotic effect.

When we were describing the facts of electrolysis on page 117, we saw that they were best explained by Faraday's conception of moving ions, so that the electrolysis of a solution of potassium chloride would be represented by the drift of a stream of positively electrified K + ions in one direction and of negatively electrified Cl − ions in the other. When we eliminate the polarization at the electrodes, Ohm's Law applies to electrolytes, that is, the most minute electromotive force produces a corresponding current. Hence it is not the electric forces which separate the ions from each other: they possess migratory freedom of themselves, and the effect of the electric force is merely to move them against the internal resistance of the liquid.

These considerations, osmotic and electrical, led Svante

Arrhenius of Stockholm in 1887 to propound a theory of ionic dissociation. According to this theory, the solution of an electrolyte contains electrified ions, parts of the molecules of the solute dissociated from each other. At great dilution, most or all of the solute is dissociated; as the concentration increases, the proportion of dissociated ions falls and that of undissociated molecules rises. Thus, in a solution of potassium chloride, we have a reversible equilibrium between KCl molecules and $\overset{+}{K}$ and $\overset{-}{Cl}$ ions $\quad KCl \rightleftarrows \overset{+}{K} + \overset{-}{Cl}$.

Later work suggests another possibility—that in freely conducting solutions all the ions are always dissociated, but that, with increasing concentration, the electric forces between them become so powerful that they reduce the effective freedom of the ions, and it is this which diminishes the osmotic pressure and the electrical conductivity.

It should be noted that the freedom of the ions is a freedom from each other. It does not imply that they are free from combination with the solvent, in fact the slow speed with which an ion moves under an electric force indicates that it is a body of considerable size, and probably consists of a central charged nucleus surrounded by an envelope of solvent, moving with it and held perhaps by the electric force. The temperature coefficient also is about equal to that of the viscosity of the solvent, and the viscosity measures the friction on a body of considerable size moving through the liquid.

With more complex solutes we may get more ions; thus the electrical relations indicate that molecules of sulphuric acid or barium chloride dissociate into three ions,

$$H_2SO_4 \rightleftarrows \overset{+}{H} + \overset{+}{H} + \overset{-}{SO_4} \quad \text{and} \quad BaCl_2 \rightleftarrows \overset{+}{Ba} + \overset{-}{Cl} + \overset{-}{Cl}.$$

This result is confirmed by the osmotic effects. The depression of freezing point of dilute solutions of these substances is roughly three times the normal.

Electrical and osmotic measurements indicate that, as the concentration of a salt in water increases, either the number of free ions or the effectiveness of their freedom diminishes. Hence water, the molecules of which are close together, should itself be almost undissociated. But even the purest water obtain-

able shows a slight conductivity, and there is reason to believe that this is not all due to impurities; there is a very slight ionization of water itself, though almost vanishingly small. This argument about concentration must not be generalized too rashly. Some pure fused salts are good electrolytic conductors and therefore are probably partly dissociated.

The almost complete non-dissociation of water has already been used in the theory of neutralization (p. 18). We saw that the heat evolved by the action of strong acids on strong bases was independent of their nature and really represented the heat of combination of the ions $H+$ and $OH-$ to form water.

$$\overset{+}{H} + \overset{-}{Cl} + \overset{+}{Na} + \overset{-}{OH} \rightleftarrows \overset{+}{Na} + \overset{-}{Cl} + H_2O,$$

or

$$\overset{+}{H} + \overset{-}{NO_3} + \overset{+}{K} + \overset{-}{OH} \rightleftarrows \overset{+}{K} + \overset{-}{NO_3} + H_2O.$$

Not all solutions of salts, alkalies, or acids are good electrolytes. Acetic acid and ammonia, for example, are very soluble in water but they conduct badly. As suggested by the theory, their osmotic pressure effects are very little more than the normal—again indicating only slight dissociation. In the equation

$$HA \rightleftarrows \overset{+}{H} + \overset{-}{A},$$

equilibrium is reached when the concentration of undissociated acetic acid HA is large, and that of the ions $H+$ and $A-$ very small.

Acids such as acetic are called weak; their acidic properties are less intense than those of acids such as sulphuric or hydrochloric. On the dissociation theory, this is clearly connected with their low conductivity. In the ordinary chemical reactions of solutions, it is the ions that are active. The tests for a chloride are tests for the chlorine ion $Cl-$, and the tests for copper those for the ion $Cu+$. If two dilute solutions be mixed, there is, in general, no change, since almost all the solutes exist as dissociated ions; dilute solutions of potassium chloride and sodium nitrate

$$\overset{+}{K} + \overset{-}{Cl} + \overset{+}{Na} + \overset{-}{NO_3}$$

when mixed do not react. But, if one possible substance is but slightly ionized, the equilibrium is upset. This is what

we found in the case of neutralization. Similarly, if a solution of sodium acetate be treated with hydrochloric acid

$$\overset{+}{\mathrm{Na}} + \overset{-}{\mathrm{A}} + \overset{+}{\mathrm{H}} + \overset{-}{\mathrm{Cl}} = \overset{+}{\mathrm{Na}} + \overset{-}{\mathrm{Cl}} + \mathrm{HA},$$

the ions H+ and A− unite to form some undissociated molecules of acetic acid. Now acetic acid is only in equilibrium with hydrogen ions when the concentration of HA is large and that of the H+ ions very small. Hence a large amount of HA is formed and practically the result of the reaction is to give sodium chloride as ions and acetic acid as undissociated molecules. This reaction used to be regarded as the "turning out of a weak acid by a strong one." It is here represented as a case of equilibrium.

The concentration of hydrogen ions is a measure of the acidity of a solution as it stands. On the other hand, neutralization with a standard alkali introduces a new substance and upsets the equilibrium till all the acid is destroyed. Hence the value of a strong acid as determined by titration is the same as that of a weak acid. To estimate the real strength of an acid as it exists, the hydrogen-ion concentration must be found electrically by measuring the electrolytic conductivity or the electromotive force of a cell which contains a standard calomel electrode (mercury and mercurous chloride in a normal solution of potassium chloride), and a hydrogen electrode (platinum saturated with hydrogen) placed in the solution to be examined.

Pure water at 21° C. has a hydrogen-ion concentration of 10^{-7} gram-equivalent per litre. This is conveniently expressed by writing $p_{\mathrm{H}} = 7$. If p_{H} is greater than 7 (i.e. if the hydrogen-ion concentration is less than 10^{-7}) the liquid is alkaline, if p_{H} is less than 7 it is acid. The values of the hydrogen-ion concentration are of especial importance in biochemistry.

Chemical Equilibrium

The reactions between molecules and ions serve to illustrate the phenomena of chemical equilibrium. In general, any chemical action between two substances *AB* and *CD*, goes on till a state of equilibrium is reached. If the reverse action can occur, and *AD* and *CB* react partially to form some *AB*

and *CD*, the final state may or may not be the same. A simple case is well shown by the reaction of acids and alcohols. For instance, by acting on ethyl alcohol, $C_2H_2.OH$, with acetic acid, $HC_2H_3O_2$, we get ethyl acetate and water, and conversely, ethyl acetate and water give alcohol and acid. At ordinary temperatures, these actions occur very slowly, but they may be hastened by heating. When they go on at the same rate in opposite directions, equilibrium is attained.

$$C_2H_5.OH + H.C_2H_3O_2 \rightleftarrows C_2H_5.C_2H_3O_2 + H.OH.$$

Another instance, in which one of the products is a gas, may be given. If chalk (calcium carbonate or $CaCO_3$) be heated, we get quicklime, that is, the oxide of calcium, CaO, and gaseous carbon dioxide. For each temperature, there exists a definite equilibrium pressure of the gas and in a closed vessel the reaction

$$CaCO_3 \rightleftarrows CaO + CO_2$$

is reversible, and analogous to the physical equilibrium between water and its saturated vapour. If, on the other hand, the gas is allowed to escape when formed, as in the ordinary lime kiln, no equilibrium is reached and the whole of the chalk is decomposed.

The phenomena of chemical reaction and equilibrium are well illustrated by Fig. 63, which represents the results of Pélabon's experiments on hydrogen and selenium. These substances when heated combine to form a gas hydrogen selenide, H_2Se, analogous to sulphureted hydrogen, H_2S.

At 270°, hydrogen and selenium unite till 0·048 of the whole is combined. This is represented by the point *B* on the diagram. On the other hand, if we begin with hydrogen selenide, at 270° it partially decomposes, but when 0·16 of the whole is left, as at the point *E* on the diagram, the reaction ceases. Thus the point of equilibrium is not the same if approached from opposite sides. Similar measurements at other temperatures lead to the curves *ABC* and *DEC* in the diagram. Above 350° these two curves coalesce, and we get the theoretical single curve *CFG* of true reversible equilibrium. Between the two curves below 350° we get states which may be described as those of false equilibrium.

All this is analogous to the equilibrium of mechanical systems. The false chemical equilibrium is analogous to the state of a body held up by friction on an inclined plane. True reversible chemical equilibrium is analogous to the mechanical equilibrium of a spring supporting a weight (p. 37), in which movement one way or the other is caused by the slightest increase or decrease in the weight.

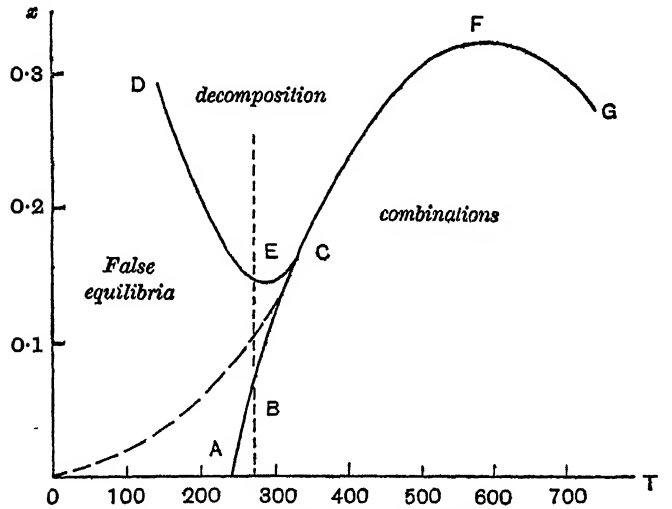


Fig. 68. Hydrogen and selenium.

Another more familiar but more complicated case is given by the relations between hydrogen and oxygen on one side and water vapour on the other.

At ordinary temperatures, hydrogen and oxygen may be mixed without combination, and, on the other hand, water vapour shows no signs of dissociation. If the mixed gases be slowly heated, gentle combination begins at 180° C. and increases steadily as the temperature rises to about 875°, when the whole has become water vapour. If the mixed gases are locally heated above this temperature by a spark, combination occurs and, owing to the heat developed by the chemical action, spreads through the mixture explosively. At about

1000° water vapour begins to dissociate, and, as the vapour is heated further, the dissociation increases in a true reversible monovariant equilibrium resembling that between a liquid and its saturated vapour.

Mass Action

Further information about chemical and electrolytic equilibrium may be obtained by considering the active mass of the reagents and products, that is, the mass per unit volume concerned in the reaction.

If single molecules are dissociating, whether into electrified ions or into other parts, the number dissociating per second will be proportional to the whole number m left, and may therefore be put equal to km, where k is a constant.

If the ions or other parts recombine, they can only do so when they collide with each other, and the frequency of collision will be proportional to the product of the numbers of each kind of ion or other reagent in unit volume. If the original molecules split into two ions or parts, the number n of each of these will be the same, and the rate of recombination will be $k'n^2$.

Equilibrium is reached when the rates of dissociation and recombination are equal, or when

$$km = k'n^2.$$

For instance, if one gram-molecule of electrolyte be dissolved in a volume V of solution, a certain fraction α is ionized, and the remaining fraction $1 - \alpha$ left as undissociated molecules. Hence the equation of equilibrium is

$$k\left(\frac{1-\alpha}{V}\right) = k'\frac{\alpha^2}{V^2}$$

or

$$\frac{\alpha^2}{V(1-\alpha)} = K.$$

This equation, known as Ostwald's dilution law, should give the relation between the ionization of an electrolyte and the dilution of its solution. $K\,(= k/k')$ should be constant at all concentrations, when α is determined by measuring the electrolytic conductivity. This is found to be in accordance with

facts for electrolytes such as acetic acid, that is, for them the equation holds good, but it does not hold for strongly ionized electrolytes, where the powerful electric forces of the many ions probably introduce disturbing factors.

Similar illustrations of mass action are found in the dissociation of diatomic molecules of iodine vapour into monatomic molecules

$$I_2 \rightleftarrows 2I,$$

and of nitrogen peroxide gas into NO_2

$$N_2O_4 \rightleftarrows 2NO_2,$$

which, at higher temperatures, itself breaks up

$$2NO_2 \rightleftarrows 2NO + O_2.$$

Again, in the Deacon process of chlorine manufacture, oxygen is acted on by hydrogen chloride

$$O_2 + 4HCl \rightleftarrows 2H_2O + 2Cl_2.$$

Where, as in these reactions, a change in the number of molecules occurs, the reaction may be investigated by measuring the gaseous or vapour pressure. In this way all these cases have been shown to conform to the laws of mass action.

The law has also been found to hold, approximately at any rate, for some reactions between liquids, but here it is difficult to know quite what is meant by concentration or active mass, so that the verification of the law is not easy.

Rates of Reaction

The general theory of the mass law of chemical action may also be applied to trace the rate at which a non-reversible action goes on.

If single molecules are breaking up with no question of interaction with other molecules, as for instance in the two first equations given above read from left to right, we have what is called a mono-molecular reaction. The number of molecules which break up per second is simply proportional to the whole number of molecules present, which, of course, gets less as the reaction proceeds. Therefore a curve drawn between time and concentration, in which the ordinates give the mass

of undissociated molecules left, must show a continual fall. The rate of reaction is given by the slope of the curve, and this must be proportional from point to point to the mass of substance left, that is, to the height of the ordinate at the point (Fig. 64). The rate of reaction, or the slope of the curve, must be equal to $k(a-x)$, where a is the original concentration, x the number of molecules per unit volume which have dissociated and k a constant.

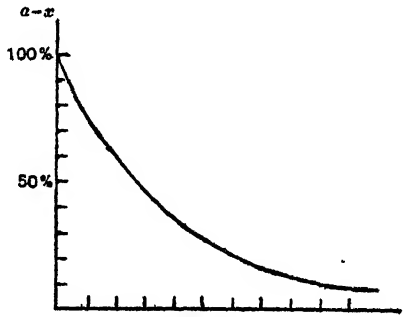


Fig. 64.

To verify this relation it is better to integrate the equation

$$\frac{dx}{dt} = k(a-x).$$

Those to whom this process is familiar will see at once that

$$k = \frac{1}{t}\log\frac{a}{a-x},$$

and this quantity should be constant if the theory is correct.

It will be seen that the ordinate $a-x$ of the curve falls in a geometrical progression as the time increases arithmetically. For instance, in Fig. 64 the ordinate falls to half value in three units of time. In the second three units this half value is diminished to its half, and so on.

It is difficult to trace such an action with a gaseous molecule, and the best examples are found in the processes of radioactivity, which we shall deal with in chapter VIII. Exact instances of the theoretical curve are there very numerous.

As we have already said, two molecules can only react together when they are in collision. If hydrogen and chlorine are combining to form hydrogen chloride in weak daylight, the number of collisions with chlorine suffered by any one hydrogen molecule will be proportional to the number of chlorine molecules present. The total number of collisions

will also depend on the number of hydrogen molecules, and will therefore be proportional to the product of the two concentrations C_1 and C_2. From all the collisions, a certain fractional number will result in combination, and hence the rate of combination will vary with the number of collisions and be equal to kC_1C_2.

If the amounts of the two reagents are equivalent,

$$C_1 = C_2 = a - x,$$

where a is the original concentration of either, and x is the number of molecules which have combined. Hence for a dimolecular reaction, the rate of reaction must be equal to $k(a-x)^2$. Those who can integrate can transform

$$\frac{dx}{dt} = k(a-x)^2 \text{ into } k = \frac{1}{t}\frac{x}{a(a-x)}.$$

Instances of dimolecular reactions are found in the two first equations of page 147 read from right to left, and also in such reactions as the "saponification" of an ester, say ethyl acetate, with an alkali (see p. 164), say caustic soda.

$$CH_3.COOC_2H_5 + NaOH \rightleftarrows CH_3.COONa + C_2H_5OH.$$

For this reaction, experiment has shown that k in the last equation is a constant. The equation has also been verified for the rate of recombination of ions in a gas which has been exposed to an ionizing agency, such as X-rays, and then removed.

More complicated reactions, in which three or four molecules are concerned, have been studied by similar methods.

Catalysis

Cane sugar reacts with water to form dextrose and laevulose

$$C_{12}H_{22}O_{11} + H_2O \rightarrow C_6H_{12}O_6 + C_6H_{12}O_6,$$

substances which will be described in chapter VI. At ordinary temperatures this reaction proceeds extremely slowly if the reagents are alone. But if a small quantity of nitric or hydrochloric acid be added, the rate of reaction is very greatly increased, though the amount of acid remains unaltered and it seems to take no part in the chemistry of the reaction. The acid is called a catalyst and its action catalysis.

Acids have the same accelerating effect on the slow spontaneous decomposition of an ester or ethereal salt in presence of water to form an acid and an alcohol

$$CH_3.COOC_2H_5 + H_2O \rightleftharpoons CH_3.COOH + C_2H_5.OH.$$

The rate of change is found to be proportional to the concentration of the hydrogen ions, being much greater with a strong acid such as nitric than with a weak acid such as acetic. It may indeed be used as a measure of the hydrogen ion concentration, giving the same results as the electrolytic conductivity or the electromotive force of a hydrogen electrode.

Another instance of catalytic action, this time with gases, is found in the combination of carbon monoxide and oxygen

$$2CO + O_2 \rightarrow 2CO_2.$$

This change does not proceed at all, or at any rate not to any measurable extent, when the gases are quite dry. The presence of water vapour, or a gas such as hydrogen chloride, that will form water vapour during the action, is necessary for the combination.

In this case there is reason to believe that nascent hydrogen is found and acts as an intermediate substance which, in contact with oxygen, re-constitutes water-vapour.

$$2CO + 2H_2O \rightarrow 2CO_2 + 4H$$
$$4H + O_2 \rightarrow 2H_2O.$$

This effect of water vapour leads to a consideration of the effect of water and other liquids when used as solvents. Many reactions will only go on in solution, and in them the solvent clearly plays a great part. It may be that in all such cases the solvent should be classed as a catalyst. It should be noted that water, containing the ions hydrogen, H, and hydroxyl, HO, is specially successful in this matter.

Whatever be the mechanism, chemical or electrical, of the process, the effect of catalysis is merely to facilitate and hasten an action. It may be compared with the effect of a lubricant such as oil applied to machinery. It seems to diminish something analogous to friction, which without it may delay movement or even hold the system in frictional equilibrium.

CHAPTER VI

ORGANIC AND BIO-CHEMISTRY

Carbon

As we saw in chapter I, the known compounds of carbon are much more numerous than those of any other element. The atoms of carbon are able to unite with each other in long chains or in rings, and to form, in combination with the atoms of other elements, an almost unlimited number of compounds, often very complex. Among these compounds are the essential constituents of the bodies of plants and animals, and hence the general study of the compounds of carbon came to be called organic chemistry. Of recent years, the chemistry of living organisms has developed into a separate subject under the name of bio-chemistry.

It used to be thought that organic compounds could be made by living things only, but many have now been prepared synthetically, that is, put together from their elements in the laboratory, and there seems no fundamental difference between the chemistry of the laboratory and that of the living cell. It does not follow that life itself, still less the consciousness which sometimes accompanies it, is a mere affair of physics and chemistry. Whether such a mechanical view of life is justified is a much deeper problem, of which the solution is not yet in sight.

The Composition of Carbon Compounds

Towards the end of the eighteenth century, Lavoisier proved that many vegetable substances were composed of carbon, hydrogen and oxygen in various proportions, while animal substances often contained nitrogen as well.

Practically all organic bodies are decomposed by red-hot copper oxide, the carbon being burnt to carbon dioxide, and the hydrogen to water. These reactions are used in analysis.

A small quantity of the compound to be examined is weighed and placed in a thin porcelain boat, narrow enough to run easily into a long tube of hard glass, which is shown in Fig. 65 lying horizontally above a row of gas burners. About half this tube is filled with granulated copper oxide. The porcelain boat is placed behind the oxide, and a roll of copper gauze behind the boat.

A slow stream of oxygen, dried and freed from all impurities, is passed through the tube from the gauze to the oxide, and the tube is heated by the gas burners. The organic compound is volatilized, and the vapour, carried on by the current of oxygen, is oxidized by the copper oxide.

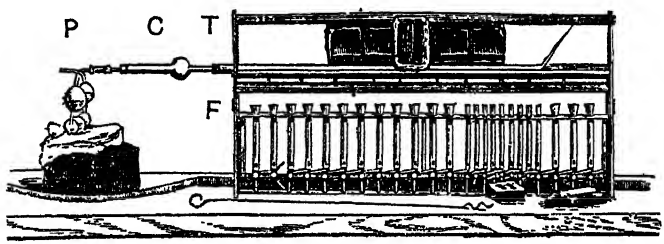


Fig. 65. Combustion apparatus.

In order to absorb the products, they are passed first through a tube *C*, filled with calcium chloride, a substance which has a strong affinity for water, and secondly through a bulb *P*, or a series of bulbs, containing a strong solution of potash, which reacts with carbon dioxide to form potassium carbonate. Any water carried on by the current of oxygen from the potash solution is caught by a second, smaller calcium chloride tube.

The first calcium chloride tube, *C*, and the potash bulb, *P*, are weighed before and after the experiment. The increase in the first gives the weight of water, and the increase in the second gives the weight of carbon dioxide formed. From these data the percentage of hydrogen and carbon in the compound can be calculated. If the compound consist of carbon, hydrogen, and oxygen only, the percentage of oxygen is at once found by difference.

For example, let us suppose that 0·1860 gram of acetic acid is weighed in the porcelain boat and a combustion carried out. The water produced weighs 0·1120 gram, and the carbon dioxide 0·2724 gram.

Of the water, H_2O, 2 parts in 18, or one-ninth, is hydrogen, that is, 0·0124 gram or

$$\frac{0{\cdot}0124}{0{\cdot}1860} \times 100 = 6{\cdot}67 \text{ per cent.}$$

of the whole compound.

Of the CO_2, 12 parts in 44 or three-elevenths are carbon, that is, 0·0743 gram or

$$\frac{0{\cdot}0743}{0{\cdot}1860} \times 100 = 39{\cdot}95 \text{ per cent.}$$

of the whole compound.

No further elements, other than oxygen, can be detected in this substance. Hence the oxygen must be

$$100 - (6{\cdot}67 + 39{\cdot}95) = 53{\cdot}38 \text{ per cent.}$$

of the whole.

Empirical Formulae

Having thus found the percentage composition, the proportional number of atoms may be deduced by dividing by the atomic weights, H = 1, C = 12 and O = 16.

$$\frac{6{\cdot}67}{1} = 6{\cdot}67, \quad \frac{39{\cdot}95}{12} = 3{\cdot}33, \quad \frac{53{\cdot}38}{16} = 3{\cdot}34,$$

numbers, which, with very small errors of experiment, are in the ratios of 2 : 1 : 1. The empirical formula of the compound is therefore H_2CO or CH_2O.

If other elements be present, special methods of analysis must be used. For instance, nitrogen may be estimated in several ways. A combustion with copper oxide in an atmosphere of carbon dioxide, gives water and nitrogen and its oxides. If these gases be passed over heated copper, the oxides of nitrogen are decomposed, and the carbon dioxide may be absorbed in potash solution. The residual gas is then nitrogen, and may be measured by volume. Again in Kjeldahl's method, now much used, nitrogenous compounds are heated with

concentrated sulphuric acid. The nitrogen is then obtained in the form of ammonium sulphate. This is decomposed by distillation from an excess of caustic soda and the ammonia is collected in a standard solution of acid and estimated by titration.

Molecular Formulae

The empirical formula of an organic compound, as determined by analysis, is not sufficient. The compound CH_2O might equally well have the composition $C_2H_4O_2$ or, indeed, $C_nH_{2n}O_n$.

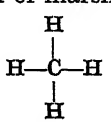
To decide between these possibilities, the molecular weight must be determined by one of the methods described on pages 12 and 45. Suppose a vapour density measurement shows that the molecular weight of acetic acid is about 60. This proves at once that the molecular formula is

$$C_2H_4O_2 = (2 \times 12) + (4 \times 1) + (2 \times 16) = 60,$$

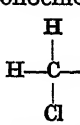
since CH_2O would give $12 + 2 + 16 = 30$, and $C_3H_6O_3$ $36 + 6 + 48 = 90$.

Constitutional Formulae

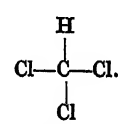
The simplest carbon compound, marsh gas, has the molecular formula CH_4. Thus one atom of carbon can unite with four of hydrogen, and the carbon is said to have a chemical valency of four. If we represent each chemical link or bond by a line, the constitution of marsh gas may be written as:



By acting on this gas with chlorine, the hydrogen atoms may be replaced successively, and we get such compounds as monochlormethane,

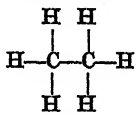


and trichlormethane or chloroform,



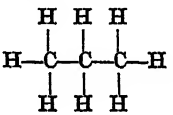
Petroleum yields a long series of similar hydrocarbons known as the paraffins.

The next simplest compound of carbon and hydrogen in this series is ethane. Its molecular formula is C_2H_6, and this clearly suggests the constitutional formula



in which the two carbon atoms, still each quadrivalent, are linked together.

Higher members of this series may be formed by again adding one atom of carbon and two of hydrogen. Thus propane, C_3H_8, is



or $CH_3.CH_2.CH_3$.

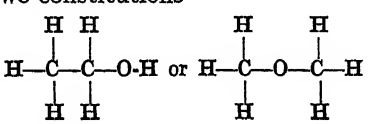
The next addition of CH_2 to form butane C_4H_{10} may be made by replacing a hydrogen atom either in a CH_3 or in a CH_2 group in propane. Hence there should be, and are, two butanes, normal butane, $CH_3.CH_2.CH_2.CH_3$, and isobutane,

or $CH_3.CH.CH_3.CH_3$,

with different physical and chemical properties. In higher members of this series of paraffins, the number of possible compounds with each empirical formula is still larger.

These bodies of identical molecular but different constitutional formulae are called isomers. The choice between possible formulae can usually be made by studying the chemical reactions of the compounds. For instance, let us trace the chemical properties of common or ethyl alcohol.

The molecular formula is C_2H_6O. This might represent either one of two constitutions



which may be written $CH_3—CH_2—OH$, or $CH_3—O—CH_3$. Ethyl alcohol is easily acted on by sodium, when one hydrogen atom is replaced by the metal, but even with more sodium none of the other hydrogens can be so replaced. Again, alcohol may be acted on by hydrogen chloride, the same atom of hydrogen and one atom of oxygen being replaced by one atom of chlorine to form C_2H_5Cl. Hence we conclude (1) that one of the hydrogen atoms in alcohol differs from the others, and (2) that this hydrogen is closely linked with an oxygen atom, since these together are replaced by one chlorine atom. Hence the constitutional formula of ethyl alcohol is clearly $CH_3—CH_2—OH$. The group OH, which is found in other alcohols, and in alkalies, is called hydroxyl.

This reasoning serves to illustrate the kind of evidence used in determining constitutional formulae.

Physical Isomers

We have seen that two butanes exist, each possessing the molecular formula C_4H_{10}, but differing from each other in chemical properties and therefore in constitutional formulae. Such bodies, as we have said, are called isomers and their existence can be explained by the principles already described, but there is another type of isomerism, known as physical or stereo-isomerism, which led to a quite new outlook on molecular structure.

Many cases are known of two bodies with the same composition, and also the same chemical properties and ordinary structural formulae, which yet are distinguished from each other by a difference in some physical property, such as the form of their crystals, their effect on a certain kind of light, their solubility, or their availability as food for higher or lower organisms.

When fused sodium chloride is heated in a colourless gas flame, it emits light of one colour, yellow, only. If a beam of this sodium light be passed through a properly prepared prism of Icelandic spar, it will traverse a second such prism placed in its path if the second prism be arranged the same way as the first one. But if the second prism be rotated about

the axis of the beam of light through a right angle, it will be found that no light can pass.

In the chapter on waves we shall see reason to believe that light is a kind of wave motion. The experiment described above is explained if we regard the first prism of spar as a kind of grating—a crystal of such structure that it only allows to pass waves oscillating in one plane. This plane-polarized light, as it is called, will clearly be able to pass through a second prism with its crystal structure placed the same way, but, if the second grating be placed across the first, the waves which pass the first will not be able to pass the second.

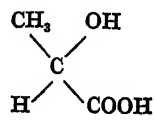
Now, if between two such prisms, arranged so that light does not pass the second, a solution of tartaric acid be inserted, the light is restored, and to get extinction again, the second prism has to be turned through a measured angle to the right. It seems then that tartaric acid has the power of rotating the plane of polarization of the light to the right—it is optically active and dextro-rotatory.

But another acid, obtained in the purification of tartaric acid prepared from grape juice, and called racemic acid, has exactly the same formula and chemical properties as tartaric acid, but is optically inactive, and rotates the plane of polarization not at all.

The sodium ammonium salts resemble these two acids: sodium ammonium tartrate is dextro-rotatory, sodium ammonium racemate is neutral. In 1848 Pasteur made a remarkable discovery. He found that the crystals deposited from a solution of the salt of racemic acid were of two kinds, identically the same except that the form of one was related to the form of the other as an image in a mirror is to the object. He picked out a number of each kind of crystal from the mixture and separated them. He then prepared with them two solutions, and found one, dextro-rotatory, to be identical with a solution of tartaric acid, and one laevo-rotatory, turning the plane of polarization to the left to the same extent as the other turned it to the right.

An explanation of these facts was given in 1874 independently by Le Bel, and Van't Hoff. They pointed out that all optically active substances possess molecules in which one

carbon atom is directly united with four different atoms or groups of atoms. Thus lactic acid, present in sour milk, has the structural formula in which each valency of the carbon atom is satisfied differently. As long as we represent this molecule by a formula written on one plane, no variation in it is possible, but there is no reason to suppose that the atoms in molecules are arranged in a plane—we have to picture them in three-dimensional space. Let us suppose that the carbon atom lies at the centre of a regular tetrahedron, connected with the other atoms at the four corners as suggested in the sketch of a solid model shown in Fig. 66. If we inter-



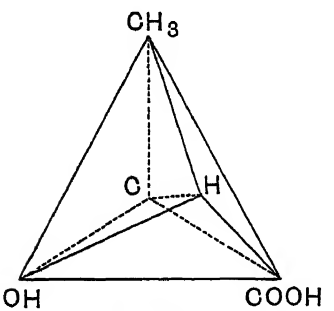


Fig. 66.

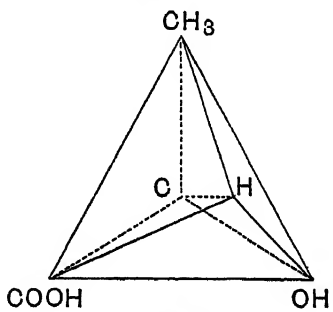


Fig. 67.

change the OH and COOH groups, as in Fig. 67, we get a picture of a different molecule, the two being related to each other as are an object and its image in a mirror. Thus in this theory it is supposed that the structure of the crystals reflects the arrangement of the molecule.

When more than one asymmetric carbon atom are present, more than two alternatives are possible. By these considerations many complex phenomena have been elucidated.

If a compound containing an asymmetric carbon atom is made in the laboratory, the product is always optically inactive, and consists, like racemic acid, of a mixture in equal quantities of the two optically active isomers. But substances occurring in nature often show optical activity, and, though chemically identical, optical isomers may be physiologically

very different. Pasteur found that the blue mould, *Penicillium glaucum*, grown on racemic acid, would use up all the dextro-rotatory isomer, leaving the other. The difference is, however, usually one of degree. The mould is able to use the laevo-modification, but only at a very much slower rate. Again, we shall see later that lactic acid is a product of muscular activity: when it is prepared from muscle, dextro-lactic acid alone is obtained. The organism is able to twist the molecules to suit itself, and is not compelled to make both isomers unless it is useful for it to do so.

Series of Organic Compounds

(1) *Paraffins.* In the various ways we have described, many organic compounds occurring naturally have been investigated and their constitution determined. From them other compounds have been prepared in the laboratory, so that a vast number are known. The majority fall into series, like that of the paraffins already described on page 155, in which each member differs from the one next below it by the same definite increment of atoms. Thus methane + CH_2 = ethane, ethane + CH_2 = propane, and any member of the series has a constitution represented by the general formula C_nH_{2n+2}. Physical properties, such as the boiling point, reflect the chemical constitution of these compounds, and also form a series: thus the gas methane, CH_4, can only be liquefied by great pressure or intense cold; normal butane, C_4H_{10} liquefies at 0° C., and normal decane, $C_{10}H_{22}$, is a solid at ordinary temperatures and does not boil till 173° C.

The paraffins occur naturally in various parts of the world in petroleum or oil-bearing strata. Their origin is sometimes attributed to the decomposition of animal or vegetable matter, or to the action of water on metallic carbides, formed by the union of carbon and metals at the high temperatures which exist deep down in the earth. Thus, aluminium carbide reacts with water to give methane and aluminium oxide:

$$Al_4C_3 + 6H_2O = 3CH_4 + 2Al_2O_3.$$

The lower paraffins are all inflammable and some of them are employed in petrol and oil-engines, since their vapours or their spray form an explosive mixture with air. The gas

issuing from the earth in oil-bearing districts is sometimes used directly for lighting purposes, and methane may form up to 40 per cent. of ordinary coal-gas. But though very readily burnt under the right conditions, the paraffins are otherwise stable, and are not acted upon by acids, alkalies, or oxidizing agents, and only slowly and in sunlight by chlorine, which gradually displaces the hydrogen atoms one by one.

(2) *Unsaturated hydrocarbons.* A gas with the formula C_2H_4 occurs in coal-gas. It is known as ethylene, or olefiant gas, and is the simplest member of another series of hydrocarbons containing two atoms of hydrogen less than the corresponding paraffins. The first few members of the series of the paraffins and of the ethylene series are:

methane, CH_4; ethane C_2H_6; propane C_3H_8; butane C_4H_{10};
ethylene C_2H_4; propylene C_3H_6; butylene C_4H_8.

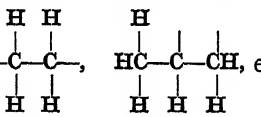
These latter compounds have the general formula C_nH_{2n} and are more reactive than the paraffins, combining directly with nascent hydrogen to give the corresponding paraffins:

$$C_2H_4 + H_2 \rightleftarrows C_2H_6.$$

They also combine directly with chlorine and bromine, hydrochloric acid, etc., and react with oxidizing agents.

$$C_2H_4 + Cl_2 = CH_2Cl.CH_2Cl.$$

This behaviour points to constitutional formulae of the type



usually written $CH_2=CH_2$ and $CH_3-CH=CH_2$, etc., indicating that the valencies of the carbon atoms are not fully satisfied, whence they are called unsaturated hydrocarbons. Even more unsaturated are acetylene C_2H_2 and its series, C_nH_{2n-2}, with constitutional formulae of the type $CH \equiv CH$, and $CH_3.C \equiv CH$. Acetylene may be prepared from its elements by passing hydrogen through an electric arc between carbon poles. From acetylene a number of organic substances may be synthetized, and can thus be built up from their elements. Acetylene is sometimes used as an illuminating gas, when it is prepared by the action of water on calcium carbide. The

latter is made by heating coke with chalk or lime in an electric furnace. The two reactions may be written:

$$CaO + 3C = CaC_2 + CO,$$
$$CaC_2 + 2H_2O = C_2H_2 + Ca(OH)_2.$$

The members of the acetylene series are very explosive, and react readily or violently with many substances, yielding saturated compounds.

(3) *Alcohols.* The simple alcohols form a series with the general formula $C_nH_{2n+1}.OH$ and may be regarded as paraffins in which one hydrogen atom is replaced by the monovalent hydroxyl group —OH.

Thus methane CH_4, ethane C_2H_6, and propane C_3H_8, give methyl alcohol $CH_3.OH$, ethyl alcohol $C_2H_5.OH$, and propyl alcohol $C_3.H_7.OH$. Since they contain the group —OH one might be inclined to regard them as organic bases. They are, however, not ionized in solution (pp. 117, 140), and have not the typical reactions of basic compounds. They are more reactive than the paraffins and undergo changes when acted upon by acids, oxidizing agents, etc., the hydroxyl group being usually affected.

Methyl-alcohol or wood-spirit, $CH_3.OH$, is obtained when wood is broken down by distillation. It is unpalatable, indeed poisonous, and is added to spirits of wine or ethyl alcohol to make it undrinkable and therefore not liable to duty. The mixture is sold as methylated spirit.

Ethyl alcohol, $CH_3.CH_2.OH$, is one of the products when sugar is attacked by the microscopic vegetable cells of the yeast plant, or by some substance produced by their activity (see page 168). The cells quickly multiply in a suitable medium, and yeast is employed to ferment starches and sugars in malt or grape juice to make various beverages, and in potatoes, molasses, etc., to produce alcohol on a large scale. Such power-alcohol, made from the quick-growing vegetation of the tropics, will perhaps be a useful source of energy to the world when the supplies of coal and oil run low.

The higher alcohols of more complicated structure include, like the corresponding paraffins, many isomers of similar empirical but different constitutional formulae. The boiling

points of the normal alcohols rise in a regular series till waxy solids such as melissyl alcohol, $C_{30}H_{61}.OH$, are reached.

Alcohols with more than one hydroxyl group also exist, such as glycol $\begin{matrix}CH_2OH\\|\\CH_2OH\end{matrix}$ and glycerine or glycerol $\begin{matrix}CH_2OH\\|\\CHOH.\\|\\CH_2OH\end{matrix}$

It may be noted that the presence of several hydroxyl groups in the molecule of a substance appears to confer on that substance the property we call sweetness: thus there are three hydroxyl groups in one glycerol molecule, five in glucose, and eight in cane-sugar. Saccharin, on the other hand, which is intensely sweet and was much used as a sugar substitute during the war, contains no hydroxyl groups, and also differs from the sugars in having no food value. Compounds containing two hydroxyl groups attached to the same carbon atom are very rare: they are usually unstable and readily lose the elements of water, the group $>C(OH)_2$ giving $>CO + H_2O$. One such compound, however, is more stable: this is chloral hydrate $CCl_3.CH\begin{matrix}OH\\\diagdown\\\diagup\\OH\end{matrix}$, prepared from ethyl alcohol and used as a soporific.

(4) *Aldehydes.* The aldehydes are characterized by the presence of the group $-C\begin{matrix}H\\\diagdown\\\diagup\\O\end{matrix}$, and may be regarded as derived from the alcohols by the removal of two hydrogen atoms. The three simplest alcohols and the corresponding aldehydes are:

Methyl alcohol CH_3OH, formaldehyde $H.CHO$;

Ethyl alcohol $CH_3.CH_2.OH$, acetaldehyde $CH_3.CHO$;

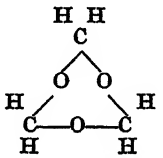
Propyl alcohol, $CH_3.CH_2.CH_2.OH$, propaldehyde, $CH_3.CH_2.CHO$.

The word "aldehyde" is a contraction of *alcohol dehydrogenatum*, a name originally given to acetaldehyde because it is formed when hydrogen is taken from alcohol. The aldehydes may be obtained by oxidizing the corresponding alcohols: thus, if methyl alcohol and air are passed over red-hot copper,

formaldehyde is produced, a pungent-smelling gas, whose solution in water is sold under the name of formalin as a disinfectant. It readily combines with oxygen to give formic acid

$$\begin{matrix} H.CH + O = H.COOH, \\ \diagdown\!\!\!\!\diagup \quad\quad \diagdown\!\!\!\!\diagup \\ O \quad\quad O \end{matrix}$$

for the carbon and oxygen atoms are united by a so-called double bond, i.e. have unsatisfied valencies, and allow the addition of one divalent or two monovalent atoms or groups. Further, molecules of formaldehyde will unite with one another, giving compounds such as trioxymethylene, $C_3H_6O_3$, which probably has the constitution represented by



It is a crystalline solid.

Acetaldehyde, $CH_3.CHO$, has similar properties, and is easily oxidized to acetic acid:

$$2CH_3.CHO + O_2 = 2CH_3.COOH + 2H_2O.$$

(5) *Acids.* The group —COOH in an organic compound confers on it the properties of an acid, the molecules being dissociated in solution, though often very slightly, to give hydrogen ions. The simplest member of the series is formic acid, H.COOH, to which the stings of ants and nettles owe part of their irritating effect. The next, acetic acid, CH_3COOH, is the essential constituent of vinegar. It is prepared on the large scale from the breakdown products of wood, and for table use from wine residues, the alcohol of which is oxidized to acetic acid by bacteria of various kinds, the chief reaction being

$$CH_3.CH_2.OH + O_2 = CH_3.COOH + H_2O.$$

Such acids form crystalline salts, e.g. sodium acetate, when they react with metallic bases:

$$CH_3.COOH + NaOH = CH_3.COONa + H_2O,$$

and esters or ethereal salts, e.g. ethyl acetate, when they react with alcohols,

$$CH_3COOH + CH_3.CH_2OH = CH_3.COOC_2H_5 + H_2O.$$

The higher acids of this series include butyric acid,

$$CH_3.CH_2.CH_2.COOH,$$

an evil-smelling compound produced in rancid butter, sour milk, or wet silage by certain micro-organisms, and also palmitic and stearic acids, $C_{15}H_{31}.COOH$ and $C_{17}H_{35}.COOH$, which form esters with glycerol of the type

$$\begin{array}{l} C_{15}H_{31}.COO.CH_2 \\ | \\ C_{15}H_{31}.COO.CH \\ | \\ C_{15}H_{31}.COO.CH_2 \end{array}$$

tri-palmitin.

Such esters are the chief constituents of natural fats, whence the acids of this series have been called fatty acids. If fats are boiled with caustic alkali, glycerol is set free and soap is formed, the latter being essentially the sodium or potassium salts of a mixture of fatty acids, e.g.

$$\begin{array}{l} C_{15}H_{31}COO.CH_2 \quad NaOH \quad CH_2OH \\ | \qquad\qquad\qquad\quad | \\ C_{15}H_{31}COO.CH \; + NaOH = CHOH \; + 3C_{15}H_{31}.COONa \\ | \qquad\qquad\qquad\quad | \\ C_{15}H_{31}COO.CH_2 \quad NaOH \quad CH_2OH \end{array}$$

tri-palmitin (fat)+caustic soda=glycerol+sodium palmitate (soap).

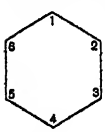
Glycerol is thus normally a by-product of soap manufacture from tallow or plant oils. The demand for nitro-glycerine (to make cordite) and the failure of imported tropical plant-oils led to a shortage of fats in Germany during the war, and a method of obtaining glycerol from sugars and hence from starch was worked out.

Benzene and its Derivatives

When coal is heated out of contact with air, it yields a multitude of gaseous, liquid and solid volatile products and a nonvolatile residue of coke. Among the liquids is coal-tar, a dark, oily substance, containing a mixture of organic compounds. When distilled, it yields a series of oils, and leaves pitch behind. Among the light oils is benzene, with a molecular

formula of C_6H_6. This substance may be prepared synthetically by heating acetylene, C_2H_2, to a dull red heat, when three molecules of acetylene combine to form one of benzene. Benzene is a very stable substance, though substitution products can be formed from it by substituting other atoms or groups for the hydrogen atoms. In these reactions the basic carbon structure of the benzene molecule seems to be unaffected. On the other hand, the formula C_6H_6 suggests an unsaturated compound. To explain these divergent indications (and for other reasons) it has been concluded that the six carbon atoms in benzene form a closed ring.

Each carbon atom is linked to two other carbon atoms and to one hydrogen atom, thus leaving one unit of valency of each carbon atom unaccounted for. Fig. 68 shows the benzene ring, but a simpler diagram may be used, the carbon and hydrogen atoms not being separately indicated.



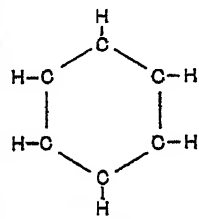
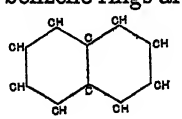
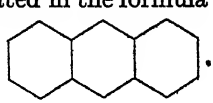


Fig. 68.

The number of derivatives of benzene is very large. When one hydrogen is replaced by another atom or group, as in chlorobenzene, $C_6H_5.Cl$, or in phenol (carbolic acid), $C_6H_5.OH$, only one compound can be formed, but, if two hydrogen atoms are replaced, there are further possibilities. First the hydrogen atoms may be next each other in the ring, in the places marked 1 and 2 in the simpler diagram for benzene given above. This arrangement forms what is called an ortho-compound. Secondly, they may be placed as 1 and 3, a meta-compound. Thirdly, if the hydrogen atoms are situated opposite each other in the benzene ring, as at 1 and 4 in the diagram, we get what is called a para-compound. It will be seen how many possibilities are opened up. In naphthalene, $C_{10}H_8$, distilled from coal tar between 170° and 230°, two benzene rings are linked together as indicated in the formula



and in anthracene three

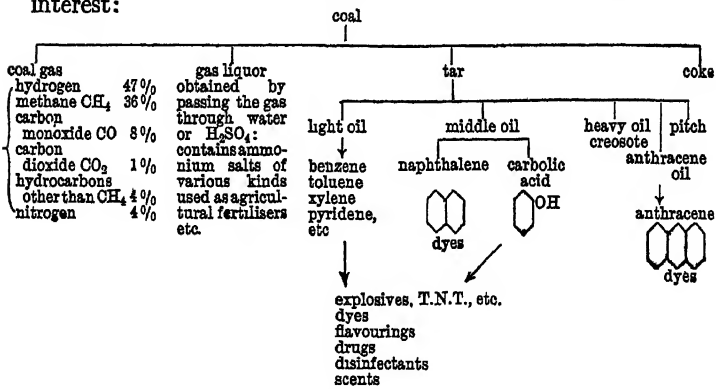


. The benzene deri-

vatives contain a large number of pleasant smelling bodies, and are hence called aromatic compounds.

Acids are formed in this aromatic series as in the fatty series by the substitution of COOH for H. Thus benzoic acid is $C_6H_5.COOH$, salicylic acid is $C_6H_4OH.COOH$, etc. Many compounds ultimately derived from coal-tar are intensely coloured or form coloured compounds with cotton, silk or wool, so that if they can be fixed in the fibre of a textile material, they may be used as dyes. For example, if they are acids or alkalies, they may combine with some constituent of the material to form an insoluble and therefore fixed salt. The chemistry of the dyes is very complicated, and its application needs both knowledge and skill.

A table of some few of the products derived from coal is of interest:



The benzene ring shows a type of behaviour common among organic compounds, namely the persistence of a definite group of atoms throughout a whole series of chemical operations. Chlorine atoms may replace hydrogen atoms round the benzene ring, or a fatty acid chain may be added, or other benzene rings may be joined on, but the original ring remains unchanged and behaves, in fact, almost like a single atom. Such persistent groups are called radicles. Examples are: $-CH_3$ methyl, $CH_3.CH_2-$ ethyl, HO- hydroxyl, = CO carbonyl, —COOH carboxyl, —CHO aldehyde, $—NH_2$ amino, $—NO_2$ nitro-groups. Radicles have definite characters of

their own, often different from those of the component atoms: thus, as mentioned above, the group —COOH is acidic, while the amino-radicle —NH_2 confers basic properties on a compound in which it is present.

BIO-CHEMISTRY

The new subject of bio-chemistry deals with the properties and characteristic changes of organic compounds as they exist in living beings. It has developed chiefly from physiology and organic chemistry, and uses the methods and results of both those sciences. But, in an increasing degree, it draws on other branches of knowledge, such as physics, bacteriology, zoology and botany. It has thus thrown new light on many problems of theoretical and of practical interest.

In living plants and animals, the chemical substances of chief importance may be grouped under three headings: carbohydrates, fats, and proteins. These substances have large molecules, built up of smaller molecular complexes, which are themselves built up from atoms. Speaking generally, the smaller complexes are made by plants, and united together in a particular pattern to form plant substances. The animal, whether eating plant or animal food, breaks down these substances and rearranges the constituent parts to form its own type of material, but does not make such parts for itself. It gets energy for its movement and other activities by breaking down some of the molecular complexes yet further into carbon dioxide and water. This energy is originally obtained from sunlight by the plant and stored in its preliminary syntheses. The linking and unlinking of the small complexes is usually a simpler affair, not involving the transfer of much energy, and rendered swift and effective by the action of enzymes or organic catalysts.

Enzymes

These little-understood bodies occur in all living tissues, and behave as catalysts, the action of which has been already described (page 150). They accelerate a reaction otherwise going on very slowly indeed, without affecting the ultimate equilibrium or themselves being altered. In fact, like other

catalysts, they seem to play a part in chemistry like that played by lubricants in inorganic machinery.

They can sometimes be extracted from the living tissues in which they exist and be made to perform their particular functions in a test-tube. This was first done in 1830, when an extract of malt (sprouting barley) was found to convert starch into sugar. The juice from yeast cells will similarly convert sugar into alcohol. Enzymes are usually less sensitive to antiseptics than living tissues, but resemble the latter in being destroyed by heat. They are most active at an optimum temperature, varying in individual cases, but usually about 45° C. Enzymes exist in tissues only in minute quantities, and have never been purified sufficiently for their chemical nature to be determined. Physically, they are colloidal. They are often remarkably specific in their action, acting only on a particular substance or on a group of substances that are similar in constitution.

An example of a type of reaction commonly catalyzed by an enzyme, is the addition or subtraction of the elements of water, as in the hydrolysis of fat. Under the influence of the enzyme lipase, present in the animal digestive system, fats are acted on by water, and give glycerol and fatty acids. The reaction takes place as shown by the following equation:

$$\begin{array}{l} C_{15}H_{31}COO.CH_2 \\ C_{15}H_{31}COO.CH \\ C_{15}H_{31}COO.CH_2 \end{array} + 3HOH \rightleftharpoons \begin{array}{l} CH_2OH \\ CHOH \\ CH_2OH \end{array} + 3C_{15}H_{31}COOH$$

Comparing this with the reaction described on page 164, it appears that lipase enables a reaction to take place in neutral solution at body temperature that otherwise needs strong boiling caustic soda to make it proceed rapidly.

Under suitable conditions, lipase also catalyzes the reverse action, namely the synthesis of fats by the removal of the elements of water from free fatty acids and glycerol. For enzymes are agents both of synthesis and of break-down: and the direction in which a reaction proceeds depends on the concentration of the reacting substances compared with that needed for equilibrium (see law of mass action, page 146) and ultimately, of course, on the changes in energy.

Reactions involving the union and separation of the smaller organic complexes are better understood than the more fundamental problem of the formation of such complexes. We will therefore deal first with the chemistry of carbohydrates, fats and proteins, and afterwards with the question of their origin.

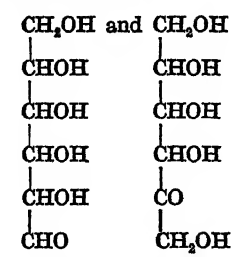
Carbohydrates

The carbohydrate group comprises sugars, starches, and celluloses, which contain carbon, hydrogen, and oxygen, the two latter elements being present in the same proportions as in water, namely two atoms of hydrogen to one of oxygen. The most important sugar for both plants and animals is glucose (grape-sugar, dextrose), a dextro-rotatory compound, soluble in water, having the molecular formula $C_6H_{12}O_6$. It is the form in which energy is supplied to muscle, and is the first product of the photo-synthetical activity of the plant. Laevulose (fructose, fruit-sugar) is a laevo-rotatory sugar with the same formula, found in fruits and honey, along with glucose. Next in order of complexity come molecules of the composition $C_{12}H_{22}O_{11}$, including cane sugar, maltose, and lactose or milk sugar, all of which are dextro-rotatory. These sugars can be hydrolyzed under the influence of hot dilute acids, or of specific enzymes, to form mixtures of the simpler sugars. Thus cane-sugar gives glucose and laevulose, maltose gives two molecules of glucose, and lactose gives glucose and another sugar called galactose, all according to the equation

$$C_{12}H_{22}O_{11} + H_2O \rightarrow C_6H_{12}O_6 + C_6H_{12}O_6.$$

In the case of cane-sugar, which is dextro-rotatory, the process of hydrolysis reverses the rotatory power, since laevulose rotates the plane of polarization more to the left than glucose does to the right. The mixture of glucose and laevulose obtained is hence called invert sugar, and the enzyme that catalyzes the reaction, invertase. This enzyme exists in the small intestine of mammals, and enables them to use cane-sugar as a food. The tissues of muscle, etc., contain no invertase, and, if cane sugar enters the circulation otherwise than by the digestive system (e.g. if injected into a vein), it is excreted untouched.

Glucose and laevulose are usually represented as having the constitutional formulae



indicating that both contain several asymmetric carbon atoms (page 158), and that a number of optical isomers can consequently exist, though few of them are known to occur naturally. The natural sugars have been studied and made from their elements through formaldehyde or through glycerol from 1890 onwards by Fischer and his co-workers, who have also synthesized many of the possible isomers.

The six-carbon sugars are the units from which the other carbohydrates are built up. Two molecules give such sugars as cane-sugar, etc., called the di-saccharides. Tri-saccharides are also known. Starch, cellulose and glycogen are poly-saccharides, containing large numbers of six-carbon sugar molecules united together; their molecular formulae can only be written as $(C_6H_{10}O_5)_n$. They are insoluble in water, and are employed by plants and animals, some as reserve materials and some as the framework of cell walls, etc.

Starch is a plant product, deposited, for example, in potato tubers, and in all kinds of grain, to supply the first needs of the new plants in spring. One can conceive the reactions involved being somewhat as follows. Six-carbon sugars are formed in the leaves and transported in the sap to the tubers or ears of corn, where the carbohydrate enzymes cause them to break down with the production of starch:

$$n\,C_6H_{12}O_6 \rightleftharpoons (C_6H_{10}O_5)_n + n\,H_2O.$$

Being practically insoluble, starch is removed from the sphere of reaction as fast as it is formed, and therefore more is continually made. When the tuber or grain is separated from

the parent plant, equilibrium is reached, a small quantity of glucose balancing a minute trace of starch in solution. So matters remain through the winter, or while the tuber or grain is stored. When supplies of moisture, warmth and light are sufficient, some kind of renewed activity causes the glucose to be used up. The equilibrium is altered, starch passes into solution and is transformed back into glucose under the action of the enzyme. As fast as it appears, the glucose is removed for the use of the germinating plant and again replaced from the stores of starch, till the latter begin to be depleted, and the green leaves of the new plant have to take over the function of providing carbohydrate.

In the germinating barley or malt used in breweries, the solution and conversion of starch go on rapidly, the first product being the di-saccharide maltose. The change from starch to maltose is catalyzed by an enzyme called diastase. The reactions in the grain are then checked by boiling, and yeast is added, which hydrolyzes maltose to glucose with the help of the enzyme maltase. The glucose is then available for alcoholic fermentation by the yeast.

Cellulose is the principal component of the cell-walls of plants, and of the framework of vegetable tissues in general. Linen, cotton and hemp fibres are mainly cellulose, and by the extraction of the other constituents with acids, almost pure cellulose may be obtained from them. It can be hydrolyzed to glucose by long boiling with sulphuric acid, at first concentrated and later dilute. By such treatment, glucose and subsequently alcohol can be prepared from wood. An enzyme which acts on cellulose has been observed in a few cases in plants, bacteria, and insect grubs, but little is known about it. It is, however, usually very difficult to demonstrate the synthetic activity of enzymes, and the breakdown of cellulose is obviously less important to plants than the breakdown of starch, as the latter is a food-store, meant to be used, while the former is of the nature of walls and a roof—i.e. a framework only. Higher animals are probably unable to digest cellulose without the help of intestinal bacteria to render it available as a food; much is excreted unchanged.

Forms of cellulose such as cotton-wool, when treated with

a mixture of sulphuric and nitric acids, give gun-cotton and other highly explosive nitrates of cellulose. Artificial silk can be manufactured by dissolving cellulose prepared from wood-pulp in an ammoniacal solution of cupric oxide, or in one of various other solvents, and forcing the solution through capillary tubes into a liquid which precipitates the cellulose in very fine fibres.

Glycogen is employed by animals as an insoluble carbohydrate reserve in the same way that starch is employed by plants. Like cellulose and starch, it has the molecular formula $((C_6H_{10}O_5)_n$ and yields glucose on hydrolysis. Some glycogen is deposited in muscles, but it is mainly stored in the liver, and can easily be prepared from liver tissue as a white amorphous, that is noncrystalline, powder. It is derived from the carbohydrates of the food, which are made soluble by the digestive enzymes, and absorbed from the intestine into the blood. Any direct requirements of the muscles and other tissues are satisfied, and the surplus is turned into glycogen for future use, unless too great an excess leads to a temporary excretion of carbohydrate, or unless a derangement of function, like that existing in diabetes, makes the organism incapable of either using or storing more than a small proportion of the carbohydrate eaten.

Carbohydrate food is the main source of energy in animals: it is used up in muscular contraction, e.g. of the involuntary muscles of the heart and alimentary canal, or of the respiratory muscles, or of muscles doing external work in digging, bricklaying, running, or making beds. The process of combustion of glucose in a contracting muscle takes place in two stages, the actual contraction involving a change of glucose into lactic acid by molecular rearrangement:

$$C_6H_{12}O_6 \rightleftarrows 2C_3H_6O_3.$$

In this process potential energy, corresponding to the work done, is lost. If an isolated muscle is made to contract by electrical stimulus in an atmosphere free from oxygen, lactic acid accumulates, and after a time the muscle becomes fatigued and ceases to respond to stimulus. In the removal of lactic acid and the restoration of the muscle to its original

condition, oxygen is consumed and carbon dioxide liberated, and it has been found that a proportion of the lactic acid is oxidized, thereby affording energy. Some of this energy is wasted in heat and some gives back potential energy to the fatigued muscle, rebuilding glucose and enabling it to do work again.

Fats

Fats are mixtures of the glycerides of the higher fatty acids (see page 164) and are neutral bodies, insoluble in water, with a low specific gravity. Natural fats and oils usually consist chiefly of the glycerides of stearic and palmitic acids, which are solid at ordinary temperatures, mixed with the liquid glyceride of oleic acid, $C_{17}H_{33}.COOH$. This is an unsaturated compound, having two incompletely satisfied carbon atoms in the middle of the long chain. The fat of milk contains also the glyceride of butyric acid, $C_3H_7.COOH$.

In seeds and nuts of various kinds, fats are deposited as reserve materials for the young germinating plant: such fatty seeds are commoner in tropical than in temperate vegetation. Palm, coconut, linseed, cotton-seed, olive and castor oils may be mentioned as familiar vegetable fats. They are used commercially for the manufacture of margarine and other butter substitutes, of soap, and sometimes of candles. Particular plant oils are employed for lighting, heating and lubrication, and for paints and varnishes, while the residue or cake from which the oil has been expressed often constitutes a valuable feeding-stuff for cattle.

When an animal eats fat, the molecules are hydrolyzed into fatty acid and glycerol, and pass the intestinal wall as these substances, but they are immediately recombined into neutral fats, which circulate in the blood. A proportion may be deposited in the tissues, often just under the skin, to form a reserve. The remainder is oxidized to provide energy. The heat of combustion (per gram-molecule) of a fat is higher than that of a carbohydrate, as the carbon is present in a less oxidized form; fats are therefore a concentrated form of reserve material.

Though vegetable and animal fats are usually included in food, animals are able to make fats from carbohydrates if

necessary, as was first proved by the experiments of Lawes and Gilbert in 1852 on the fattening of young pigs on barley. Much more fat was formed than could have come from the amounts of fat (and protein) that were consumed. The reverse change, from fat to carbohydrate, is known to occur in the germination of fatty seeds, when starch and cellulose may be formed from the fat. The fat or carbohydrate undergoing the change is probably divided into small two-carbon units for rearrangement. Though carbohydrate may replace fat to a very large extent in the food of higher animals, the reverse is not true, and a definite quantity of carbohydrate is necessary, for it is involved in the final stages of the oxidation of fats. In the absence of carbohydrate, or in diabetes when the carbohydrate processes are upset, fats are not completely burnt to carbon dioxide and water, but remain in a half-oxidized condition in the form of aceto-acetic and β-oxy-butyric acids, $CH_3.CO.CH_2.COOH$, and $CH_3.CHOH.CH_2.COOH$, which disturb the neutrality regulation of the tissues, and eventually poison the nervous system. These acids are also produced when an animal suffers from starvation.

Regulation of Neutrality

The regulation of neutrality in the tissues of higher animals deserves mention as a most interesting example of the delicate adjustments which exist in the chemistry of the living cell. It is found that the important factor biochemically is not the absolute amount of acid or alkali present, but the concentration of hydrogen or hydroxyl ions. Pure water contains $\frac{1}{10000000}$ gram of H ions in a litre. This is a perfectly neutral solution, since the same number of OH ions are also present. $\frac{1}{10000000}$ may be written 10^{-7}, and, as explained on page 143, for convenience the index or power of 10 is employed, without the minus sign to express the hydrogen-ion concentration, and is termed the p_H value. The p_H of pure water is 7. Raising the number of hydrogen ions lowers the value of p_H, while increasing the number of hydroxyl ions raises the value of p_H. The usual p_H of human blood is 7·4, rather on the alkaline side of neutrality. Blood, and other physiological fluids, contain large amounts of so-called buffer substances, that is,

substances able to prevent much change in p_H on the addition of acid or alkali. Sodium bicarbonate is such a buffer substance. When carbon dioxide is dissolved in water, the resulting carbonic acid is ionized to a small extent to give bicarbonate and hydrogen ions:

$$H_2CO_3 \rightleftharpoons \overset{+}{H} + \overset{-}{HCO_3}.$$

But let us imagine bicarbonate ions to be already present from the ionization of sodium bicarbonate

$$NaHCO_3 \rightleftharpoons \overset{+}{Na} + \overset{-}{HCO_3}.$$

The ionization of carbonic acid is now checked by these bicarbonate ions, which are already present nearly in the amount required for equilibrium. As a necessary consequence, there is very little ionization of the carbonic acid when added, fewer hydrogen ions are set free and less alteration in p_H occurs. Bicarbonates, phosphates and proteins in the blood all act as buffers, and the hydrogen ion concentration normally remains very nearly constant. Suppose lactic acid to be entering the blood as a result of muscular activity: there is a slight increase of hydrogen ions in spite of the action of the buffers. But the increase of p_H is instantly noted unconsciously by the part of the brain which regulates the respiration, deeper and more frequent breaths are taken automatically, and carbon dioxide is removed more quickly than usual through the lungs. There is consequently a decrease in the number of bicarbonate ions, a necessary decrease in the number of hydrogen ions, and the momentary increase of p_H is over. As an outward and visible sign, the man who has been taking violent exercise is seen to pant.

Proteins

Under the name of proteins is classed a large group of nitrogenous organic materials, including many substances of great importance in the chemistry of living bodies. For the most part, their chemical constitution is very complicated. But, in recent times some advance towards its interpretation has been made.

By the insertion of the amino-group —NH_2 into a fatty acid molecule, a compound is produced known as an amino-acid, which is at the same time both basic and acidic in character. The simplest example is glycine or amino-acetic acid, $H_2C.COOH$.

NH_2

About twenty different amino-acids are known. In each case the amino group is attached to the same carbon atom as the —COOH group. These acids are the main building stones out of which the proteins are constructed. All the naturally occurring amino-acids are optically active, with the exception of glycine, which has no asymmetric carbon atom. Of the others, some are dextro-rotatory and others are laevo-rotatory.

White of egg, lean meat, the casein of milk, are familiar bodies mainly of protein substance. All proteins contain carbon, hydrogen, oxygen, nitrogen and sulphur, and some, e.g. casein, include phosphorus. The molecules probably have a minimum molecular weight of about 5000, and are extraordinarily complex. Any one protein of definite origin, when purified as far as possible, has constant properties and is a well-characterized chemical compound. The casein of, say, cow's milk, will differ by some subtle factor from the casein of goat's milk, and if it were possible to observe it, one might even find minute differences between the milk or muscle or brain proteins of different individuals of the same species. It is known that human beings can be divided into groups with regard to blood characteristics, and if blood is to be transfused from one man to another, it is essential to discover a donor belonging to a group suitable to supply the patient. A question of protein constitution may possibly be partly involved here, and also in the intricate matters of immunity from bacterial infection, and even in the phenomena of heredity.

It is easy to realize how such vast numbers of proteins can exist when we remember the possible permutations and combinations of the twenty amino-acids which have been found as constituents of proteins. They are linked together like the acid and base in a salt, the amino-group of one being

joined to the carboxyl group of another, with the elimination of water. Thus, for example, with two molecules of glycine:

$$CH_2NH_2.CO\vdots OH + H\,NH.CH_2.COOH \longrightarrow$$
$$CH_2NH_2.CO.NH.CH_2.COOH + HOH.$$

The group —CO.NH— is called the peptide linkage. It is clear that such linkages can proceed indefinitely with the amino-acids, both ends of the di-glycine molecule being available for a further combination. Syntheses on these lines were carried out by Fischer, who succeeded in making in the laboratory a substance containing 15 glycine and 3 leucine (that is, isobutyl-amino-acetic acid) molecules, having a molecular weight of over 1200, and in some ways resembling a natural protein.

By boiling with acids or alkalies, or by the aid of enzymes, the peptide linkage can be undone, and when proteins are thus hydrolyzed, a mixture of amino-acids is the product. A protein may not include all the amino-acids, but a considerable number of them, in varying proportions, are present in all the proteins that have yet been examined. The analysis of amino-acid mixtures presents great difficulties, as the reagents employed are apt to destroy some of the molecules. About 85 per cent. of the weight of some proteins has been accounted for in the form of amino-acids.

When an animal takes protein food, the digestive enzymes break down the proteins into amino-acids, which are absorbed from the intestine into the blood. Those needed for protein-repair and growth are taken up by the appropriate tissues, and the rest help to satisfy energy requirements, after the removal of ammonia. In the liver, ammonium carbonate and other salts are converted into urea, in which form the surplus nitrogen of man and some other animals is excreted. The formula of urea is usually written as $(NH_2)_2CO$, but it has lately been shown more probably to have the constitution $HN{=}C\left\langle\begin{matrix}NH_3\\ |\\ O\end{matrix}\right.$. Urea is of interest as the first organic compound to be made in the laboratory (1828).

It is probable that the animal organism can, if necessary, make for itself the simpler amino-acids, such as glycine, from carbohydrates or from fats and ammonia, but the more complex ones, including all those with a benzene nucleus, have to be derived from other animals or plants, and growth cannot take place, nor life be maintained for long, without a supply of them. For example, the chief protein of maize, called zein, is deficient in one of these, and on a diet of carbohydrate, fat and zein, mice were found by Hopkins to live for about 20 days only. Proteins as such are not necessary to maintain animal life, since a mixture of amino-acids obtained by the hydrolysis of proteins with enzymes or acids is adequate to replace protein in a diet.

Much work has been done to discover what are the essentials of diet for man and the higher animals. A minimum of protein is required for the growth and repair of tissues, and of carbohydrate for the complete oxidation of fats (see page 174). But if enough of both is supplied beyond these minima, will the diet be satisfactory? For a long time this was thought to be the case. But when specially purified carbohydrates, fats, and proteins, along with the necessary inorganic salts and water, were given to rats, something was found to be missing. This "something" has been traced down to three accessory food factors (or vitamins), namely:

(1) one existing in green plants and in animal fats such as codliver oil and butter, whose absence prevents normal growth and may be one of the causes of the disease called rickets;

(2) one existing in fresh vegetables and fruits, especially turnips and lemons, the absence of which causes scurvy;

(3) one existing in the outer layers of various kinds of grain, in meat and eggs, which protects animals from a disease of the digestive and nervous systems called polyneuritis, or (in man) beri-beri.

The chemical nature of these factors is unknown, and they exist in minute quantities only. Though their absence has such disastrous effects, their presence in a diet in more than the necessary minimum quantity seems to have no effect. They may perhaps tentatively be compared with enzymes.

The Synthetic Action of Plants

For the synthesis of amino-acids containing the benzene ring, for the provision of carbohydrates, and finally, for the formation of these accessory food factors, the animal world is dependent on the green plant. Even if a man eats beef, the beef comes from grass. So the reactions occurring in the plant are of special interest to us. Yet very little is known of them. In the presence of the green pigment chlorophyll, the plant uses energy from sunlight to unite carbon dioxide and water, setting free oxygen and producing sugar.

$$6CO_2 + 6H_2O \rightarrow C_6H_{12}O_6 + 6O_2.$$

The intermediate steps have been much discussed but not clearly elucidated. Formaldehyde, HCHO, has often been suggested as the first product, and it is true that formaldehyde will easily unite with itself to give compounds of high molecular weight of either chain or ring structure. We may have as a first step,

$$2CO_2 + 2H_2O \rightarrow 2HCHO + O_2,$$

and then the union of six molecules of formaldehyde to make sugar,

$$6HCHO \rightarrow C_6H_{12}O_6.$$

Yet this is at present only speculation.

All the plant's carbon is derived from the small percentage of CO_2 normally present in the atmosphere (·03 per cent.). Both animals and plants give off CO_2 as an end-product of their chemical changes, and the air would ultimately become unfit for animal respiration if it were not for the activity of plants in removing carbon dioxide. A plant breathes out very much less CO_2 than it takes up, for, though the absorption of CO_2 only goes on in sunlight, the plant's respiration is small: it does not use energy in movement of all kinds like an animal. The plant stores the energy of light radiation in the form of potential chemical energy: the animal dissipates this chemical energy as heat, and in the performance of external work.

One certain thing about the photo-synthetical activity of the plant is the necessity of the pigment chlorophyll. This is

a green substance of very complicated chemical structure, containing a body called pyrrole, an atom of magnesium and an alcohol called phytol. The pigment of the blood, haemoglobin, also contains pyrrole nuclei, with iron instead of magnesium, united to a protein called globin. Haemoglobin has the important function of combining with oxygen in the lungs and liberating the oxygen again in the tissues where it is required. In some lower animals (e.g. crabs, etc.) the blood pigment is blue and contains copper instead of iron. But the chemical resemblance between chlorophyll and haemoglobin is usually considered to be an interesting coincidence only.

In order to build up plant proteins, the first products of photo-synthesis have to be united with ammonia to give amino-acids. The plant obtains all its nitrogen from the soil, usually in the form of nitrates, though ammonium salts can be used: the change from nitrates to ammonia is another of the important changes that have not been followed in detail. Soil nitrogen comes from three main sources. It comes (1) from the decaying bodies of plants and animals, which are broken down to carbon dioxide, water and ammonia by bacteria. Other bacteria oxidize ammonia to nitrites and nitrates, e.g.

$$(NH_4)_2CO_3 + 3O_2 \rightarrow 2HNO_2 + CO_2 + 3H_2O.$$

Again nitrogen is obtained (2) from the action of soil bacteria which can use the actual free gaseous nitrogen of the air as their source of nitrogen. And finally (3) nitrogen in agricultural practice is supplied as ammonium salts or nitrates added to the soil as such.

Ammonium salts may be obtained as a by-product of the distillation of coal. Nitrates come in large quantities from deposits in Chile, or from atmospheric nitrogen by the new process of Haber, or by the union of nitrogen and oxygen in the electric arc (see pages 17 and 18).

Conclusion

The carbon and nitrogen cycles of plant and animal life have now been briefly indicated. Other necessary substances should be mentioned.

Besides the obvious requirement of oxygen for respiration, certain mineral salts are found essential. Sulphur, entering into the constitution of proteins, and playing a part in cell oxidations, resembles nitrogen; for it is absorbed by the plant in the oxidized form of sulphates, which have to be reduced for incorporation into proteins. The sulphur in decaying proteins is oxidized back to sulphate by bacteria. Phosphorus, as phosphate, is probably important in the changes in carbohydrates, and may have to be added artificially by the farmer. Potassium, calcium, sodium, magnesium, iron and manganese are always found in plants and, in smaller or larger quantities, are necessary for growth.

Finally, water is a necessity of all living cells: all the chemical reactions of life go on in solution in water, and one is reminded of the probable aquatic origin of living beings. Some thinkers even go so far as to suggest that the concentration of the various salts in vertebrate blood may represent the concentration of the same salts in the sea at the period when the first amphibia left the water for the land.

Those who wish to learn more about the changes in living matter should read the companion volume to this one—Sir Arthur Shipley's *Life*.

CHAPTER VII

WAVES

Wave Motion

THE principles of wave motion may be illustrated by some simple experiments. A number of balls are hung up by threads and connected together by spiral springs (Fig. 69). The first ball is struck: it moves towards the next, compresses the spring, and thus starts a motion which travels from ball to ball till the last one moves outwards. If that ball be connected to the frame of the apparatus by a spring, the motion

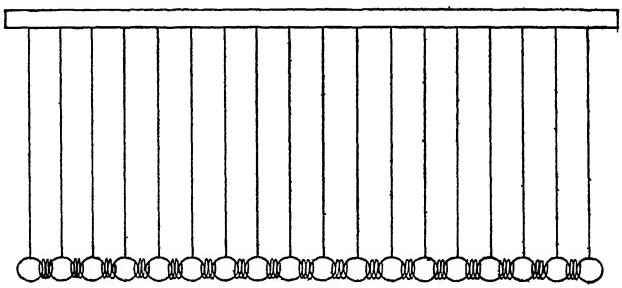


Fig. 69.

of compression will be reflected backwards till it reaches the first ball again. On the other hand, if the last ball be free, it swings out further than the others, and pulls the next ball after it; thus a wave of compression is reflected as a wave of extension. During all this motion, each individual ball moves only backwards and forwards near its position of equilibrium, but something travels from one end of the system to the other. This something is called a wave, in this case a longitudinal wave, because the movement of the balls is along the direction in which the wave itself travels.

As a second illustration, one end of a long cord of india-rubber is fixed to a wall, and the other end held in the hand.

The cord is then stretched. If the hand be jerked from side to side or up and down, a hump will be started and will run along the cord, till it is reflected back from the further end. This shows a transverse wave, in which the motion of the vibrating particle is at right angles to the direction of the wave motion.

It is easy to demonstrate that, the stronger the springs in the first illustration and the more the force exerted in stretching in the second, that is, the greater the elasticity of either system, the faster the wave moves. On the other hand, if the balls be made heavier, or a thicker rubber cord be chosen, that is, if the mass be increased, the slower is the wave. Thus the essentials for the propagation of a wave are elasticity and inertia, and on the ratio between them the velocity of the wave depends.

Whether the wave be longitudinal or transverse, each vibrating particle oscillates about a position of equilibrium. In either case, we can draw a diagram in which the horizontal axis *OT* represents either time or else distance in the direction of the wave motion, and the vertical axis *OD* the displacement of an oscillating particle from its position of equilibrium whether that displacement be along or across the direction of the wave.

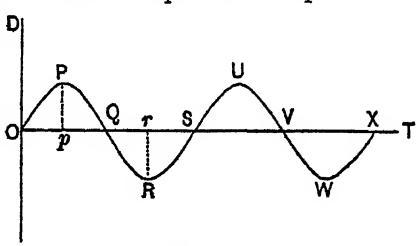


Fig. 70.

If a particle originally at *O* be given an impulse, its inertia leads to an increasing displacement. Owing to the opposing force due to the elasticity, the displacement does not increase without limit, but reaches a maximum, represented by *pP*, and then diminishes again. At *P* the velocity is nil, and at *Q*, as at *O*, the displacement is zero, and the velocity is a maximum. The particle, owing to its inertia, will then swing in the reverse direction till its displacement is *rR*, and again, under the effect of the elasticity, return towards its original position as represented by *S*.

If *OT*, instead of the time, represents the direction in which this wave is travelling, the wavy line *OPQRSUVWX* gives the displacement of successive particles in that direction instead of the successive displacements of one particle as time goes on.

If the displacement varies as does the sine of a steadily increasing angle, that is, as *CM* varies in Figure 10 on page 38, we get what is called a simple harmonic curve, and the movement of any particle is called simple harmonic motion.

One complete wave, including both to and fro motion, is represented by the line *OPQRS*. The length *OS* is called the wave length, usually written as λ; the maximum displacement, *pP* or *rR*, is called the amplitude *a* of the wave, and the time occupied by the oscillation from *O* to *S* the period of vibration *T*. The number of periods in one second is called the frequency *n*, and the number of waves in unit length the wave number. A wave, starting from *O* with a velocity *v*, traverses *n* wave lengths, such as *OS*, in one second. Hence the whole distance traversed in one second, that is the velocity, is

$$v = n\lambda. \qquad \text{.....(1)}$$

Canal Waves on Water

A water wave in a channel may be studied by putting water in a long glass trough and giving it a push forward by means of a flat piece of wood at one end. A wave is thus

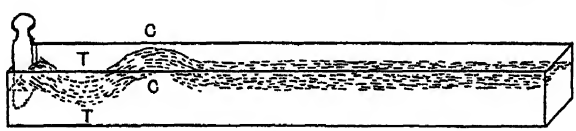


Fig. 71.

started, travels along the trough with a definite velocity, and is reflected at the farther end.

When the wood is moved, the water in front of it is heaped up. Its weight makes it fall again, and in falling push out the water in front of it, thus carrying on the wave. Owing to inertia the water falls through its position of equilibrium, and thus a wave trough *T* follows the wave crest *C*.

But surface tension also comes into play. The surface of water may be regarded as covered by a stretched, elastic film (see p. 49) and any disturbance, by increasing the area of the surface, sets up an opposing elastic force and starts a wave.

If the waves are very large, they are propagated by gravity, if very small, by surface tension. In the latter case they are called ripples. At intermediate values, both causes are concerned.

With the glass trough it may be shown that the longer the wave length of a gravity wave, the faster it moves. On the other hand, the velocity of a surface ripple is less the greater the wave length. Thus there is a certain wave length at which the wave travels more slowly than do those which are either longer or shorter. This may be taken as the boundary between true waves and ripples.

The actual motion of the particles can be traced by putting some bran or floats in the water. We can see that the water swings to and fro horizontally more than vertically. The waves are called long waves.

Again, with two similar troughs, one containing twice as much water as the other, it can be shown that the velocity of a trough wave or canal wave is proportional to the square root of the depth of the water. On a canal 4 or 5 feet deep the velocity of the wave is about 12 feet per second, or 8 miles per hour, and a boat propelled at this speed is opposed by much less resistance than at speeds rather less or more—it rides along on the crest of the wave.

If, as usual, the boat moves more slowly than the natural long wave, it sets up forced waves, and the first wave, spreading out from the bow of the boat, must adjust its wave length so that its velocity is equal to that of the boat.

Sea Waves

In deep water, waves travel with a velocity which may perhaps depend on gravity, on the wave length of the wave, and on the density of the water; that is, in terms of a dimensional equation (see p. 52),

$$v \propto g^x \lambda^y d^z,$$

$$LT^{-1} = (LT^{-2})^x L^y (ML^{-3})^z.$$

Equating the indices of M, T and L, we get:

from M $\quad z = 0$;

from T $\quad -2x = -1$, therefore $x = \frac{1}{2}$;

from L $\quad x + y - 3z = \frac{1}{2} + y = 1$, therefore $y = \frac{1}{2}$.

Hence it follows that the velocity of a wave does not depend on the density of the water but only on gravity and the wave length. Hence

$$v \propto \sqrt{g\lambda}.$$

Mathematical investigation gives $v = \sqrt{\frac{g\lambda}{2\pi}}$, which can be expressed in the simple rule that the velocity reckoned in miles per hour is equal to the square root of $2\frac{1}{4}$ times the wave length measured in feet. Thus waves which are 100 feet from crest to crest move with a speed of $\sqrt{225} = 15$ miles an hour. Atlantic waves sometimes are as much as 500 feet in length, with a period of 10 seconds and a speed of about 35 miles an hour.

With deep-sea waves, the vertical motion of the particles is more important than in canal waves, and, by watching the motion of floating corks, they may be seen to move in nearly circular paths. The amplitude of these paths decreases very rapidly as we go beneath the surface. At a depth of a few hundred feet the worst storm leaves the ocean quite untroubled.

When waves reach shallow water, the friction of the water against the bottom retards the backward part of the circular motion of the water particles, while no such effect is found above on the forward motion. Moreover, the front of the wave is in shallower water than the back of the wave, and hence the back part travels faster and continually overtakes the front. For both reasons, the forward side of the wave gets steeper and steeper, till the roller eventually curls over and breaks on the shore.

The production of waves, of course, uses up energy, and thus the bow and stern waves of a ship oppose resistance to its motion. The total resistance to a moving ship is due to three causes: (1) waves, (2) eddies in the water, (3) skin friction. The skin resistance is diminished by using a hull

with a smooth surface, cleaned periodically. The other resistances depend on the form of the ship and its speed. Experiments on the resistance offered to ships are now made in experimental tanks, some as much as 600 feet long. A movable bridge crosses the tank, and can be run along it by motors on rails laid down on each side of the tank. Models of a proposed ship are towed by the bridge, and the force exerted on them is measured at different speeds. Thus the best form for the ship and the engine power needed may be estimated, and vessel and engines designed accordingly.

Sound Waves

If a gong which can be sounded continuously by clockwork be hung by silk strings inside a glass bell jar and the jar be exhausted by an air pump, the sound diminishes, and, at a moderate exhaustion of the air, ceases to be heard. Hence with this arrangement air is necessary for the propagation of sound.

When a sounding gong is touched with a cork ball hanging by a thread, the ball is violently jerked about, showing that the gong is vibrating. It is clear that a vibrating surface will alternately compress and rarify the air in contact with it, and the evidence of these two experiments together indicates that the sound which affects our ears consists of longitudinal waves of alternate compression and rarefaction travelling through the air—longitudinal because the vibration of the particles of air is in the direction in which the sound travels.

In the simple illustrations of wave motion given above, we saw that the speed of the wave increased when the elasticity increased and decreased with greater mass (p. 183). And, in the general case, Newton proved that the velocity of any wave travelling through a medium was given by the expression

$$v = \sqrt{\frac{E}{d}},$$

where E is the particular elastic constant involved, and d the density of the medium. We saw on p. 48 that the isothermal volume-elasticity of air is measured by its pressure, and the density of air at 0° C. is 0·001293 gram per c.c. Hence at

standard atmospheric pressure (p. 46) the velocity should be

$$v = \sqrt{\frac{1{\cdot}014 \times 10^6}{0{\cdot}001293}} = 2{\cdot}80 \times 10^4$$
$$= 280 \text{ metres a second.}$$

Now, by timing the interval between the flash of a distant gun and the arrival of the sound, the velocity has been shown to be about 331 metres a second. The cause of this discrepancy was pointed out by Laplace, the great French mathematician and astronomer. When air is compressed, work is done, and consequently the air is heated. If the compression is slow, the heat leaks away and the air remains at one temperature. Its elasticity is then that determined by Boyle's law, and measured by the pressure. But the alternate compressions and expansions of a sound wave occur very rapidly, perhaps several hundred times a second—too rapidly for the heat to escape. In this case the elasticity is greater, and it may be proved that it is greater in the ratio γ of the specific heat at constant pressure to that at constant volume (see p. 69). Hence the equation becomes

$$v = \sqrt{\frac{\gamma p}{d}},$$

and, since for air γ is 1·41, $v = 332$ metres a second, in agreement with the observed value.

In other gases the velocity of sound can be calculated from their densities and specific heats. For instance, since the density of hydrogen is only the 1/14th that of air and γ the same, the velocity of sound in hydrogen will be $\sqrt{14} = 3{\cdot}7$ times its velocity in air. (See also p. 192.)

In liquids and solids the volume elasticity is very great; hence the velocity of sound is high. Thus in water the elasticity is $2{\cdot}04 \times 10^{10}$ C.G.S. units and the density unity. Hence v comes out as 1428 metres a second, and has been found by experiment to be 1435 metres.

A sound travelling through air is much affected by atmospheric conditions such as wind. Suppose that a sound wave is moving against the wind. The wind is retarded by friction near the ground, and hence moves quicker above. The upper

part of the sound wave is therefore retarded more than the bottom, and the wave front gets tilted backward. The sound moves at right angles to the wave front, and thus the sound passes upward into the air and may become inaudible on the ground. Sometimes it meets other currents of air above, and may be refracted downwards again. Hence there is sometimes a zone of silence, beyond which the sound is once more heard.

Differences of temperature produce similar effects. At the same pressure the density of hot air is less than that of cold air; hence a rise in temperature increases the velocity of sound. A sound wave, therefore, is refracted by differences of temperature in the medium, the sound tending to bend in the direction of the colder regions. These considerations show the difficulties which exist in the propagation of sound signals, such as the sirens used by lighthouses in foggy weather.

During the war, the accurate comparison of the time of arrival at several different stations of the sound of a gun, was developed into an exact art of "sound-ranging," the position of an invisible gun being thus discovered. The sound was recorded automatically and corrections were applied for such of the above described causes of disturbance as could be estimated.

Music

If the vibrations of the source of sound are regular, the ear recognizes the sound as pleasant or musical. If, on the other hand, the source is subject to an irregular disturbance, the sound is mere noise. The liking for regular, rhythmic movement is, of course, not confined to the ear, we feel it also in riding, rowing and dancing.

Any body of which the natural frequency of vibration lies between about 40 and 5000 per second may be used as a musical instrument. The piano and violin depend on the oscillation of wires or strings, the organ and cornet on the vibration of columns of air. The full sound of each kind of instrument depends on a principle called resonance.

If a child's swing be given a series of gentle pushes in exact time with its natural period, the vibrations will increase more and more. On the other hand, if the pushes are not properly timed, it will never "get up a swing."

If a tuning fork, say one which gives 256 vibrations a second, be struck, it emits a faint sound. If it be held over the top of a tall glass cylinder and water be poured in, it will be found that, at one particular length of the column of air, about 1·1 feet, the sound becomes much louder than it is for any other length. The natural period of oscillation of this air column is the same as that of the fork; it therefore absorbs the energy emitted by the fork, is set into vibration and resounds, though for a shorter time than the fork sounds if alone.

In an organ pipe, air is blown through a mouth-piece partially closed. A number of sounds are started: from them all, the organ pipe picks out that one with which it is in unison, and gives it out as a loud resounding note.

On the other hand, in the piano or violin it is the wire or string with a definite period of vibration which starts the sound. The wooden parts of the instrument can vibrate to any pitch, and thus reinforce the particular sound of the string, whatever it may be.

The vibrations of strings can be illustrated by the rubber cord described on page 182. By a transverse jerk of the hand, a wave can be started. After reflexion at the far end, this returns to the hand and may then be reinforced by another jerk. The cord is thus set vibrating transversely, the ends being at rest or nearly so, and the middle part oscillating backwards and forwards. By giving twice as many jerks per second, the cord can be made to vibrate in two halves, another fixed point or node appearing at the middle, and two

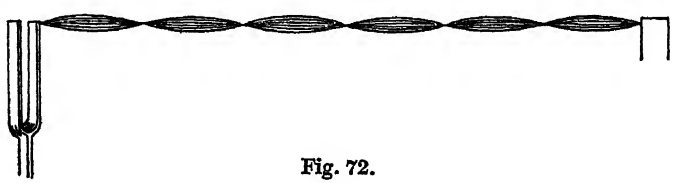


Fig. 72.

loops or internodes being formed on each side. By even quicker jerks the cord may be broken into three loops with four nodes, two at each end and two between, and still more nodes and loops may be formed by a string attached to a tuning-fork (see Fig. 72). These nodes and loops are formed

by the interference of direct and reflected waves. The waves appear to stand still on the cord or string and are hence called stationary undulations. They may also be seen when water waves are reflected from a solid wall.

These observations can be correlated with musical sounds by means of a wire stretched over two bridges by a weight hanging over a pulley. If this wire be plucked aside at its middle point and let go, it forms one loop, and emits a musical note. If, while vibrating, it be touched lightly at the centre, or if a bridge be placed there, it oscillates in halves and emits a note which a musical ear recognizes as the octave of the first. The velocity being the same and the wave length halved, equation (1) on page 184 shows that the frequency must be doubled. We conclude that the octave vibrates twice as fast as the fundamental note, its period of vibration being half as great.

An organ pipe must be open at the mouthpiece, but it may be open at the top or closed. Any closed end must be a node; the air is there at rest. At an open end the air can move when vibrating, and this end is therefore an internode or loop. Since the wave length is double the length between two nodes, the wave length of the fundamental note of a closed organ pipe is four times the length of the pipe. If the upper end of the pipe be open as well as the lower, the simplest mode of vibration gives one node at the centre and one internode at each end. Here the wave length is twice the length of the pipe.

When a string or an organ pipe is emitting its fundamental note, other vibrations are superposed on the simplest, and the fundamental note is accompanied by the octave and other harmonics. It is the number and relative strength of these overtones or harmonics which give its quality to the sound—which distinguish, for example, the sound of an organ from a piano, and both from that of the human voice.

Velocity of Sound in Gases

The vibration of columns of gases gives a method of measuring in them the velocity of sound. A long glass tube placed horizontally has one end tightly closed by a cork through which is thrust a glass rod to half its length. The

other end of the tube is closed by another cork which can slide within it, and thus adjust the length of the column of gas. The tube is dismounted and inclined at an angle, and through it is poured a stream of lycopodium spores—very light grains. They leave a streak of dust behind them. The tube filled with air is then placed in position again and the corks inserted. By rubbing the free end of the glass rod with a damp cloth, it can be set in longitudinal vibration, its clamped middle point being a node. The disc at its inner end, as it oscillates to and fro, throws the air into vibration,

Fig. 73.

and, by adjusting the position of the cork at the other end, a position will be found where the length of the air column is a multiple of half wave lengths and the lycopodium spores fall into groups as indicated. The average distance between two nodes, that is, half the wave length of the sound, may thus be estimated. The air is then replaced by another gas and the experiment repeated. The frequency of oscillation of the rod being the same, the velocity in the new gas is to the velocity in air as the new wave length to the old one. From what has been said on page 188, it is clear that, from a knowledge of the velocity of sound in a gas, the ratio of its specific heat at constant pressure to that at constant volume may be determined.

Interference of Sounds

Two notes, nearly but not quite in unison, give rise to a sound which is alternately louder and softer. These alternations are called beats. Their cause is easily seen if we investigate the joint effect of the two notes by means of a curve of displacement as given in Fig. 74.

The curve *A* gives the displacement curve of one sound, the curve *B* of the other of slightly greater frequency. The curve *C* is obtained by adding together the ordinate of the two other curves from point to point. At *d* and *f* the two curves *A* and *B* show displacements in the same direction,

so the resultant amplitude, the sum of the two, is large. At *e* the displacements of *A* and *B* are equal and opposite, so that the resultant displacement of *C* is zero. Hence at *d* and *f* we get loud sound, and at *e* silence, the number of beats per second being equal to the difference in frequency of the two notes. When these beats come quicker than about ten a second we cease to hear them separately. Their effect then is to spoil the tone of the notes, the worst discord being with notes whose difference of pitch gives about thirty beats a second.

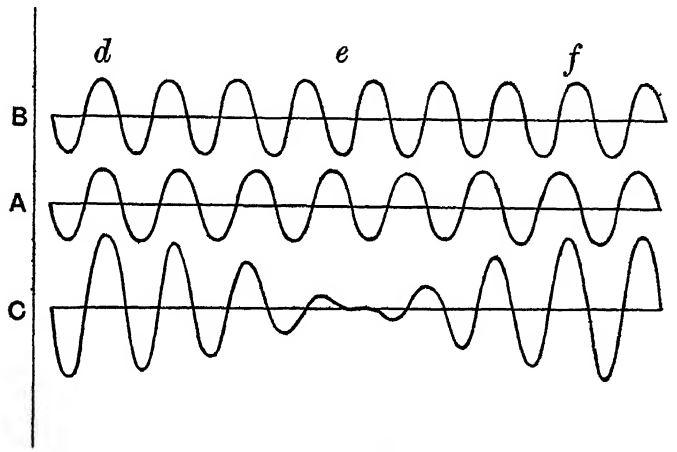


Fig. 74.

Another effect of interference is to produce sound shadows behind a large obstacle like the shoulder of a hill. A small obstacle throws no sound shadow; sound bends round it and is audible on the other side, but, when the obstacle is large compared with the wave length of the sound, the area of silence is quite marked; in this case, sound travels in straight lines, and we may almost talk of rays of sound as we do of rays of light.

To get some idea of the reason of this effect, let us study Figure 75. Let *S* be a source of sound, and *O* an observer. Let *MPM'* represent the sound wave-front at a certain moment

after it has left S. At that moment the sound energy exists as a wave-pulse on this front, and any final effect at O must somehow be produced by this wave-front.

In considering beats, we saw that two equal waves in opposite phases, as at e in Figure 74, cancel each other and produce silence. This occurs when one wave is just half a wave length in front of the other, when, in fact, a crest of one coincides with the trough of the next. Thus, if one of two equal waves starts half a period before the other, their joint effect is nil.

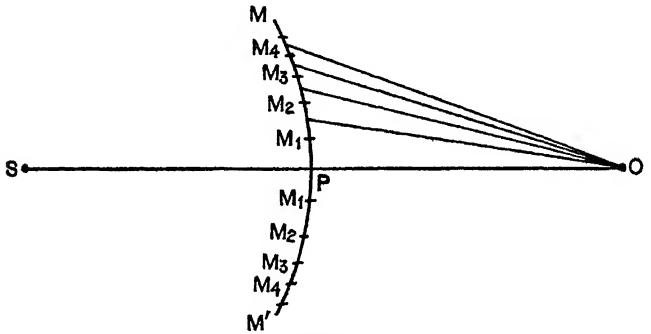


Fig. 75.

Now divide the wave front MPM' in Figure 75 into a series of rings round P as centre in such a way that the first ring M_1PM_1 is on the whole just half a wave length nearer O than the next ring M_2M_1 M_1M_2, and so on. If straight lines be drawn as shown from O to P and from O to the middle of each successive element, OP is one half wave length shorter than the next line, and so on.

It can be shown mathematically that, for small wave lengths, while M_1M_1 produces at O a greater effect than the opposite effect produced by M_2M_1 M_1M_2, as we pass away from P the effects of the elements soon become equal as well as opposite. Each element of the wave-front at MPM' then destroys the effect of the next element when the wave reaches O, and the only resultant effect at O is due to a very few elements just round P, the point of the wave-front which lies

on the straight line from S to O—in fact, the only sound audible at O has travelled along or very near to the straight path SPO, and if a large obstacle be interposed at P so as to cut off these effective elements, there will be silence at O.

This result is only attained when a large number of half-period elements can be drawn on the trace of the wave-front, that is, when the wave length is short. If the wave length be long, it is obvious that the number of half-period elements will be small, and they will all produce an effect; a sound shadow is only formed when the number is large, so that, when passing away from P on the wave-front, we soon come to the regions where their effects are equal as well as opposite. Hence we see that sound shadows are formed only by obstacles which are large compared with the wave length of the sound.

Interference of Light

The phenomena of interference, described above in the case of sound, are of even more interest and importance when they appear with light. Indeed, it is interference which shows that light is an instance of wave motion.

A simple way of showing the interference of two rays of light, due to Young, is to allow a beam of sunlight to enter a dark room through a slit S and then pass through two small pin holes or slits A and B very near together. On a screen at M, light from the two sources A and B will overlap, and a series of coloured bands are seen. If, instead of sunlight, the yellow rays from a gas flame containing sodium are used, a series of bright and dark lines appear at M.

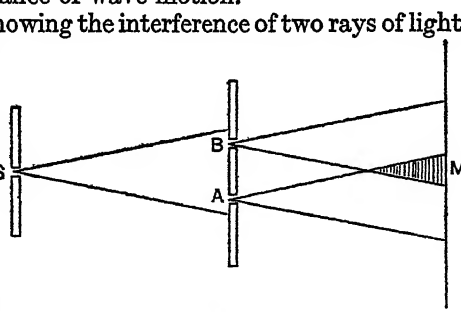


Fig. 76.

These interference lines are at once explained by a wave theory of light. At the middle point of M the distance from A and B is the same, and the two waves, originally from the

same source S, and therefore identical in character, reinforce each other. But a little way from M we come to a point where the distance from A is half a wave length greater than that from B; the two waves differ by half a period, cancel each other's effects and give darkness instead of light.

Similar effects were obtained by Fresnel with two prisms or with two mirrors inclined at a very small angle, so as to be not quite in the same plane. Two images of a single source are thus obtained very near together, which correspond to the apertures A and B in Figure 76, and produce the same effects.

The colours of thin films, such as soap bubbles or the layer of air between two plates of glass pressed together, are also due to interference. They are best examined by placing a convex lens in contact with a flat glass plate or one face of a prism. A series of coloured rings—known as Newton's rings —are then seen. If yellow sodium light is used, the rings are alternately bright and dark.

The explanation depends on the fact that when one half of a train of waves is half a wave length in front or behind the other half, the troughs of one train coincide with the crests of the other, interference results and there is darkness. When light falls on the upper surface of the film of air, part is reflected and part transmitted. Of the transmitted part, some again is reflected at the lower surface and emerges parallel with that reflected from the upper surface, having travelled over a path longer by the distance covered in air. When these two rays of light differ in phase by half a period there is darkness, and when they differ by a complete period there is light. At the upper surface, where light leaves the denser medium, reflexion itself causes a change in phase of half a period, as explained on page 182, where the wave of compression travelling through a system of balls and springs, is reflected as a wave of extension. Thus the centre of the

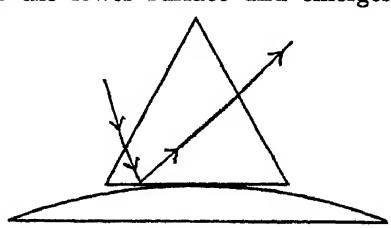


Fig. 77.

ring system, where this change of phase only occurs, is dark—as indeed must be the case since the glasses are in contact and there is no reflected light. As we look away from the centres, we pass places alternately where the two rays are in the same or opposite phases and thus see a system of light and dark rings.

Remarkable interference effects are obtained by means of a diffraction grating—a smooth reflecting or transparent surface on which parallel scratches are ruled many thousands to the inch. When a beam of parallel light from a slit falls on a transparent grating, the scratches stop the light, but each smooth interval $a'b'c'$, etc., in Figure 78, becomes the centre of a disturbance on the further side. Usually the wavelets spreading out from these elements interfere with each other and no light is seen, but, at one particular angle, they arrive all in the same phase. These parallel rays reinforce each other when brought to a focus by a lens, and a bright line, an image of the slit, is seen, where each path differs from the next by one complete wave length λ, i.e. where

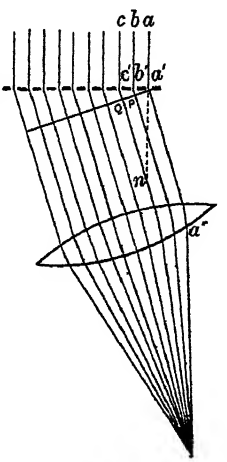


Fig. 78.

$$\lambda = b'P = a'b' \sin b'a'P.$$

Thus, $$\lambda = d \sin \theta, \qquad \ldots\ldots(2)$$

where d is the distance between the intervals or the scratches on the grating, and θ is $b'a'P$, or $na'a''$ the angle between the incident light and the direction in which a bright line appears. The grating is mounted on the table of a spectroscope (page 217), and parallel light from a sodium flame admitted. The eye-piece is moved till the exact position of the bright line is brought into coincidence with its cross wires. Thus the angle θ is measured, and λ, the wave length of the light, calculated from the formula.

A multitude of other interference phenomena are known, and it is their evidence which has established the wave theory of light.

The chief difficulty in the way of that theory, a difficulty which led Newton himself to represent light as due to a number of minute particles or corpuscles shot off from the luminous object, was the fact that light travels in straight lines and throws sharp shadows of obstacles placed in its path. But the scale of interference effects shows that the wave lengths of light are very small indeed, and measurement with a grating gives about 0·0006 of a millimetre, the one fifty-thousandth part of an inch. Hence the reasoning given to explain the existence of somewhat imperfect sound shadows shows that with light very sharp shadows are to be expected. This is confirmed by the theory of the grating, for, if the scratches were not there, those elements would send forward waves of opposite phase, the bright line at the angle θ would be destroyed, and the incident light travel on in a straight path. Hence, speaking generally, the shadows thrown by light are sharp. Nevertheless, it is quite easy to show by experiment that, if light from a pin-hole source passes the straight edge of an obstacle, interference fringes are seen just outside the boundary of the shadow, and that, inside the shadow, the light does not stop quite suddenly, but fades away continuously though rapidly, in exact accordance with the requirements of the wave theory. Usually our sources of light are not point sources, and the interference effects from different parts of the source are superposed and invisible. Moreover, owing to the minuteness of the wave length, the interference effects are on so small a scale that the shadows seem very nearly sharp. It follows, then, that, in accordance with the history of the subject, we may, if we please, develop a geometrical science of optics on the idea of rectilinear rays of light.

Reflexion of Light Waves

On the whole, it is more interesting and amusing to treat light as waves, and trace the effect on them of reflexion and refraction*. As in Figure 79, let a spherical wave diverge from a source O and fall on a plane reflecting surface AB. At some definite instant, let $ANQB$ represent the position

* See Preston's *Theory of Light*.

the spherical wave would have reached if there were no reflecting mirror. As things are, the wave-front meets the surface at A and B. Each point of the mirror between A and B will have reflected the part of the wave which falls upon it, and thus each point will have become the source of a reflected wavelet. For instance, the middle point M will be the centre of a spherical wavelet, which has reached N', a point as far in front of the mirror as N, the point which the original wave would have reached had no mirror intervened, is behind it. Any other point P is the centre of a reflected wavelet of radius $PQ' = PQ$.

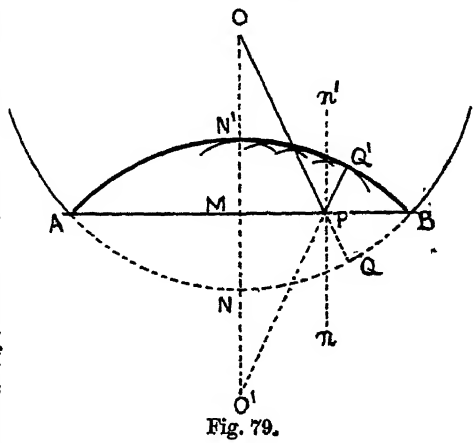


Fig. 79.

Now each of these reflected wavelets would produce an infinitesimal effect, indeed, at any arbitrary point, they interfere with each other. But they all touch the arc of a circle $AN'B$, and form a reflected wave front, for there alone they arrive all in the same phase and do not interfere. The reflected wave-front $AN'B$ is clearly an arc of a circle the centre of which is O', as far behind the mirror as the source or object O is in front. The point O' is called the image of O in the reflecting surface.

Hence it is seen that the effect of a plane reflecting surface is simply to reverse the curvature of the incident wave.

If we regard O as the source of diverging rays of light, a reflected ray PQ' will lie in the same plane as the incident ray OP and the normal nn' to the surface, and the angles between the normal and the incident and reflected rays will be equal. These are the laws of reflexion established by experiment.

It is now easy to trace the path of the light by which the

eye sees an object in a mirror. Some of the rays diverging from each point of the object OO' enter the eye. They seem to come from corresponding points II'. Hence the eye sees an image II' as far behind the mirror as the object is in front. It should be noted that if O be the right hand of a man standing in front of the mirror, the corresponding point I of the image is its left hand. Hence, in a sense, the image is the reverse of the object, and the country behind the mirror is a "looking glass world."

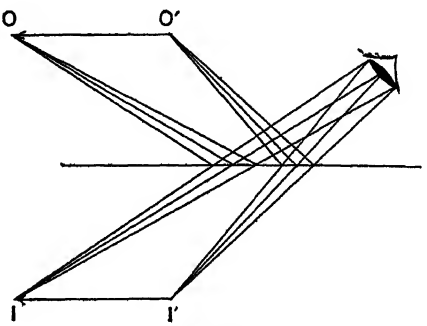


Fig. 80.

If the mirror be concave, as AMB in Figure 81, a spherical wave diverging from O, which would have reached ANB if there were no mirror, will be reflected to $AN'B$ where $MN' = MN$. Hence

$$DN + DN' = 2DM.$$

Now the curvatures of the spherical surfaces are measured by the so-called sagittae DN, DN' and DM, and so we may write

$$c_1 + c_2 = 2c,$$

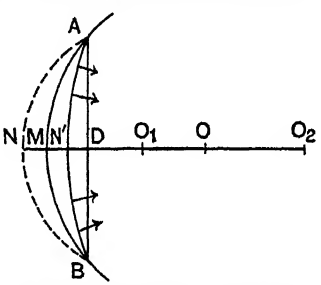


Fig. 81. Concave mirror and wave fronts.

where c_1 and c_2 are the curvatures of the incident and the reflected waves and c the curvature of the reflecting surface. Now the curvatures are inversely proportional to the radii of curvature r_1, r_2 and r. Hence the equation becomes $\frac{1}{r_1} + \frac{1}{r_2} = \frac{2}{r}$. But r_1 and r_2 are the distances u and v of the object and image from the mirror. Thus we get

$$\frac{1}{u} + \frac{1}{v} = \frac{2}{r}. \qquad \text{.....(3)}$$

If the mirror is convex, its curvature must be reckoned negative. With this convention the equation still holds.

When the object is a great way off, so that all light coming from it travels parallel to the axis, $1/u$ is zero, and $v = r/2$. Thus the image of a distant object is at a distance from the mirror equal to half its radius of curvature; its position is then known as the principal focus of the mirror.

We can now trace by the method of rays the formation of an image in a spherical mirror. Let OO' in Fig. 82 be the object. The rays OC and $O'C$, which pass through the centre

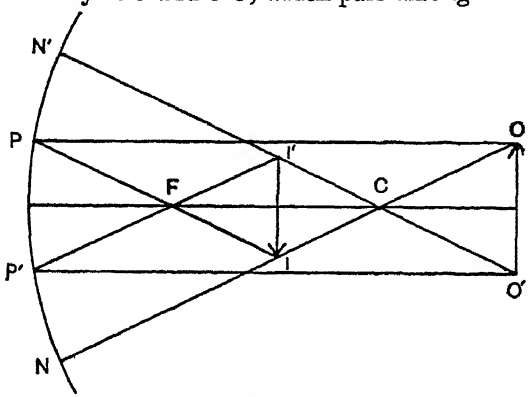


Fig. 82.

of curvature C, fall normally on the mirror at N and N', and are therefore reflected back on their own paths. The rays OP, $O'P'$, which run parallel to the axis, are reflected through F, the principal focus. The points at which these rays cut the normal rays gives the position of the image II'. Compared with the object it is inverted, and proportional in its linear size to its distance from C.

Refraction of Light Waves

The general theory of the refraction of waves at the boundary of two media was given long ago by the Dutch philosopher Huygens (1629–1695), though it was not till Fresnel showed the effect of interference that a complete account could be given.

Let AB represent the front of a wave of light incident in air on the surface AA' of a denser medium such as water or glass. In this medium the light is assumed to travel more slowly. Hence, while the light is travelling from B to A' in air, the wavelet which spreads from A as centre will travel over a shorter length—let us say over a semicircle of radius AC shorter than BA'. The wave-front AB will meet the surface in successive points between A and A', and each element of the surface will, like that round A, become a centre of disturbance. The resultant wavelets will intersect along a line CA'. Here all the wavelets will surge up together in the same phase, and produce a new wave-front in the

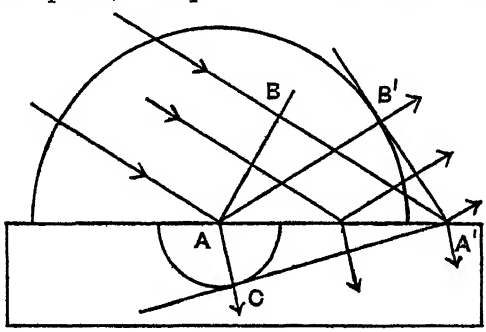


Fig. 83.

denser medium. Elsewhere the phases of the wavelets will be different; interference results, as explained on page 194, and there is no effect. Part of the energy of the incident wave will be reflected, and in the same way a reflected wave-front is formed at $A'B'$, the radius AB' of the wavelet spreading from A being equal to BA'.

Now let us return to the refracted waves. In the denser medium, since the light will move at right angles to the wave-front, the ray through A will be refracted in the direction AC. Now AC was drawn so as to bear to BA' the proportion of the velocity v' of the light in the denser medium to its velocity v in air. Hence

$$\frac{v}{v'} = \frac{BA'}{AC} = \frac{BA'/AA'}{AC/AA'} = \frac{\sin BAA'}{\sin AA'C} = \mu,$$

where μ is the constant ratio of the velocities, called the refractive index of the medium in contact with air. The angles BAA', $AA'C$, the angles between the wave-fronts and the surface, are the same as the angles between the rays and the normal to the surface, which we called the angles of incidence and refraction, i and r, respectively. The equation

$$\frac{\sin i}{\sin r} = \mu \text{ (a constant)} \qquad \text{.....(4)}$$

expresses correctly the experimental law of refraction.

Besides the light which enters the second medium, some of the incident light is reflected, and the proportion between the energy of the reflected and refracted waves depends on the angle of incidence. When falling from a rarer on to a denser medium, the energy is always divided, but light falling from the denser medium on the surface of separation may fail to get out altogether. In this case

$$\frac{\sin i}{\sin r} = \frac{1}{\mu}, \text{ and when } \sin i = \frac{1}{\mu}, \text{ then } \sin r = 1,$$

that is, the angle of refraction is 90° and the refracted ray is parallel to the surface. At and beyond this angle of incidence, then, there is no refracted ray.

Total reflexion may be seen by looking from below at the upper surface of water in a drinking glass held above the level of the eye.

It is easy to see why water looks shallower than it really is. If unobstructed, the wave spreading out from an object O at the bottom of a pool would have spread over a segment of a sphere the trace of which is ABC. But in air the refracted wave will have travelled farther, as ADC. The curvature is thus increased, and the centre of the new

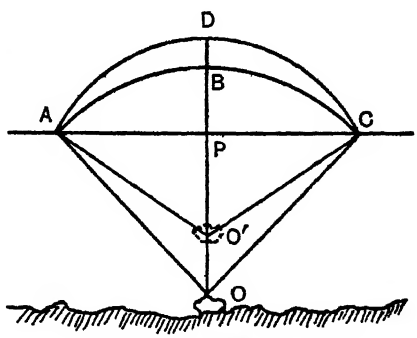


Fig. 84.

spherical wave is at O' a higher point, and from O' the light seems to diverge. O' is the image of O, and, to an eye outside the water, O seems raised to O'.

In Figure 84, $\frac{PD}{PB} = \mu$ and PD and PB are proportional to the curvatures, or inversely proportional to the radii of curvature of the waves. Hence

$$\frac{OP}{O'P} = \mu,$$

or the ratio of the real to the apparent depth is equal to the index of refraction. By focussing a microscope on O, first with and then without water between, the distance the microscope is raised gives OO'. By focussing on dust particles floating on the surface of the water, the new rise gives OP, and hence μ may be determined. Its value varies much with different media, even with different kinds of glass, and also with the colour of the light used, being greater for blue than for red. To this point we shall return.

Most optical instruments involve the use of lenses—that is, of pieces of transparent refractive substances with polished convex or concave surfaces.

A lens bends the incident light towards the thicker parts. With a convex lens, rays coming from a long way off are refracted through a point on the axis called the principal

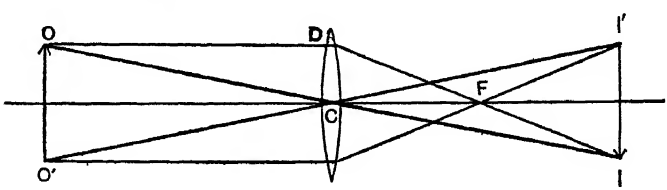


Fig. 85.

focus, F. Rays going through the centre of a lens C will be refracted as much one way on emergence as they are the other on incidence; hence, after passing the lens, they travel in a direction parallel to their original direction. These two rules enable us to trace the position of the image of an

object OO', formed by a thin lens. The ray OC proceeds undeviated; the ray OD passes through F. Where these intersect at I, the image of O is formed.

If the object is further from the lens than the focal length, as in Figure 85, the image is formed on the further side of the lens and the rays pass through it, so that it is real and can be thrown on a screen. If the object is between the lens and a principal focus, as in Figure 86, the rays ODF and OC diverge and will never form a real image; to an eye beyond the lens they seem to come from a point I on the same side of the lens as the object, and I is then called a virtual image—it cannot be thrown on to a screen.

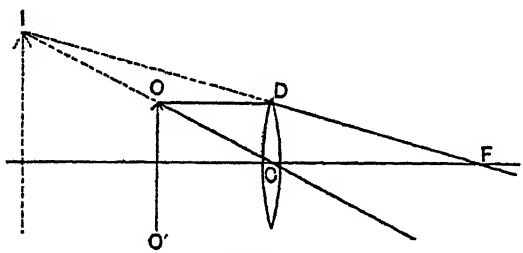


Fig. 86.

By considering the changes of curvature of the waves of light as they pass through the lens, as was done with a mirror on page 200, it may be shown that, if f be the focal length, u and v the distances of the object and image respectively from the lens,

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v}.$$

Hence, by observing the positions of the object and the image with light of any given colour, the focal length of the lens for that light may be measured. It will be found to be shorter for blue than for red. Hence with white light, which, as we shall see, is formed of a mixture of coloured lights, the image is surrounded by a fringe of colour, since all the rays cannot be brought to a focus together.

The use of lenses in optical instruments may be illustrated by the compound microscope. In Figure 87, the object glass G,

a convex lens, forms a real inverted image at *I*. This is viewed through an eye-piece *E*, placed nearer than its focal length. A virtual image *I'* is thus formed, also inverted with regard to the object, and very greatly magnified. In the figure, the paths of rays of light are shown by fine continuous lines, and dotted lines are used to indicate the imaginary paths along which rays appear to come to an eye which takes all paths of light to be straight.

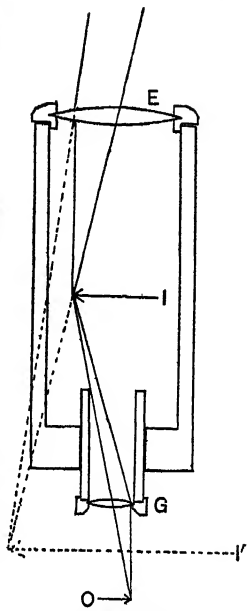


Fig. 87.

In the astronomical telescope, the convex object glass is large and forms a real inverted image of the distant object very near the principal focus. This is viewed through a small convex eye-piece, and again an inverted magnified image is formed. For terrestrial telescopes the image is again inverted so that it is erect compared with the object.

Dispersion

Newton was the first to investigate systematically the dispersion of white light into its constituents. He allowed a beam of sunlight to enter a dark room through a hole in a shutter, and placed a glass prism in its path with a screen beyond it. Instead of a white image of the hole, a band of coloured light, ranging from red through orange, yellow, green and blue to violet, was seen on the screen. All the rays were refracted towards the thicker part of the prism, but the red was found to be refracted least and the violet the most.

The difference between the refractions of light of different colours varies with different kinds of glass, and is not necessarily proportional to the refraction. Thus the dispersion colours seen round the image formed by a simple lens may be corrected. A convex lens of crown glass produces general refraction but little dispersion. A concave lens of flint glass

placed in contact with it will bring together the red and the blue rays, while leaving much of the general refraction. Hence the combination will act as a convex lens of longer focal length, and show very little, if any, dispersion colour.

Experiments on dispersion are best carried out by means of a spectroscope or spectrometer in which the prism is carried on a table which can turn on a central axis. The light is admitted through a slit of adjustable breadth placed at the principal focus of a convex lens. The light emerges from the lens as a parallel beam. It is refracted through the prism, and then enters a telescope, by which the eye sees with homogeneous light a sharp image of the slit, and with white light a coloured rainbow-like spectrum. The prism is turned on its table till the deviation of the light passing through it is least, when its path is symmetrical with respect to the prism.

If the source of light be a colourless gas flame with a small quantity of a fused sodium salt placed in it, the light is yellow, and the spectrum consists of a single yellow line, which, if two prisms be used instead of one, can be resolved into two lines. Salts of other metals give other and equally characteristic spectra, while yet other lines are obtained by the light of gases rendered incandescent by an electric discharge. This is the origin of the science of spectrum-analysis; the angular position with a given prism of the lines of many elements may be determined, and these elements then recognized by a glance at the spectrum of an unknown mixture. When this fact was understood, it was seen that the existence of spectral lines which could not be assigned to any known substance indicated the existence of undiscovered elements. In this way the alkali metals caesium and rubidium were discovered by Bunsen in 1860 and 1861.

The Physics and Chemistry of the Sun and the Stars

The white light from an arc lamp or from incandescent lime gives a continuous coloured spectrum. But if a narrow beam of sunlight be admitted into a dark room and passed through the spectroscope, the coloured band of light is seen to be crossed by a number of fine, sharp dark lines called after Fraunhofer, their discoverer. Many of these dark lines coincide

exactly in position with the bright lines from incandescent elements. This is shown on Plate I, the frontispiece.

The explanation of this remarkable fact seems first to have been given by Sir George Gabriel Stokes in his lectures at Cambridge. It is essentially a case of resonance, resembling the phenomena in sound described on page 189. The inner parts of the sun are intensely hot, and emit complete white light. The outer envelope is cooler, and, in accordance with the principle of resonance, absorbs energy of the particular frequencies of vibration it would itself emit. Thus, the gases and vapours which exist in the envelope of the sun deprive the light which is radiated into space of certain factors, and these missing factors are made evident by the dark Fraunhofer lines. The envelope of the sun is, indeed, hot and radiates, but its radiation is much less intense. Nevertheless, by covering the bright parts of the spectrum, the faint illumination of the Fraunhofer lines may be demonstrated.

This explanation was given independently and verified experimentally by Foucault and by Bunsen and Kirchhoff. In 1860 the latter observers passed the white light from incandescent lime through the cooler flame of a spirit lamp, and showed that, when a sodium salt was introduced into the spirit flame, a dark absorption band was seen in the coloured spectrum in the same position as the bright sodium line from the sodium flame itself.

The element helium may be said to have been discovered by Janssen, who, during an eclipse in 1868, observed a bright yellow line in the spectrum of the sun's chromosphere not coincident with those of any known element. This hypothetical element was afterwards given the name helium by Frankland and Lockyer, and its presence on the earth was detected by Ramsay in 1895 in the gas which is contained in certain rare minerals containing the radioactive metal uranium. The meaning of this fact will be shown later.

The immediate significance of Bunsen and Kirchhoff's discoveries lay in this: they revealed to mankind that the familiar physics and chemistry of our terrestrial laboratories held good also in the sun and stars. Newton had proved the same result for mechanics, and shook to its foundations the age-long

idea that the divine heavens were altogether and essentially different from the lowly earth. The proof that solar and stellar spectra could be interpreted in terms of terrestrial elements completed the overthrow of this older mode of thought. The heavens were no more divine than the earth, or rather the earth became as divine as the heavens. The physics and chemistry of the sun and stars lay open to the eyes of man. The spectra of many fixed stars showed them to be similar in constitution to our sun, while the line spectra of certain nebulae indicated that they were gigantic masses of glowing gas—worlds in the making.

Accurate measurement of the position of spectrum lines led to another new field of investigation. Nothing at first sight would seem more hopeless than the determination of the velocity of a star approaching or receding from the earth in the line of sight. Yet this is now a commonplace observation in solar and astro-physics. When an express train rushes whistling through a wayside station, an observer on the platform does not need a very accurate ear to note a sudden lowering of pitch in the sound as the engine passes. As the source of sound recedes, each vibration has further to travel to the ear, and thus fewer of them arrive per second, that is, the pitch is lowered. Similarly, when source and observer approach each other, more waves are crowded into the ear per second and the pitch is raised.

This is called the Doppler effect, and it applies also to light. If a luminous body is moving towards the observer, the "pitch" of the waves is raised, that is, the lines are shifted towards the blue end of the spectrum; if the body recedes, the lines move towards the red. Thus, by measuring the minute difference from the normal in the position of known lines in the spectrum of a star, or of a mass of incandescent gas in the atmosphere of the sun, its velocity towards or away from the earth may be estimated. The application of all these principles to astro-physics will be considered in chapter IX.

The Range of Spectra

There is no very simple relation between the wave length and the refraction in spectra formed with a prism; but spectra

may be produced by diffraction gratings as described on page 197, and with diffraction spectra the simple formula

$$\lambda = d \sin \theta$$

gives the connexion between the wave length λ and the angle of deviation θ, d being the distance between the scratches on the grating.

As measured in this way, the wave lengths of the visible rays of light are found to range from about 7600 tenth-metres in the red to 4000 tenth-metres in the violet, a tenth-metre being 10^{-10} metre. In considering sound, we saw that an octave comprised the notes from one of a given frequency of vibration to one with double that frequency. The ear can appreciate as musical sounds the notes of about seven octaves. But the wave lengths of the visible spectrum show that the eye has a much smaller range than the ear, indeed that the whole of the visible rays form rather less than one octave.

It is easy to show that waves exist beyond the visible spectrum at each end. Let us first consider the heating effect, which measures the total energy of the waves. Let us suppose that a spectrum of an incandescent solid is formed by a prism of fluor-spar, since glass absorbs much radiant heat, or by a reflecting diffraction grating. Then let the heating power of the spectrum be explored by means of a blackened thermo-couple or very thin strip of platinum used instead of the coil of wire in the electric resistance thermometer described on page 58.

Beginning at the violet end, we find very little result; the effect increases as we move the thermometer into the green and yellow, becomes large in the red, and goes on increasing for some little distance in the infra-red region beyond the limits of the visible light. It then falls off gradually. In Figure 88, the ordinates are proportional to the readings of the thermometer from point to point along the spectrum, and the abscissae proportional to the wave lengths, the unit being the micron μ, the one-thousandth part of a millimetre. The visible spectrum ranges from about 0·4 to 0·8. The lowest curve corresponds to a temperature of 1000° C., the middle curve to 1250°, and the top curve to 1450°. In each case the

highest point in the curve lies in the infra-red, but it shifts towards the shorter waves as the temperature rises.

On the other hand, it is possible to photograph rays considerably beyond the violet end of the visible spectrum. These ultra-violet waves have intense chemical, electrical and physiological effects. They are easily absorbed, and their presence is more marked as we ascend high mountains where the light has passed through less of the atmosphere.

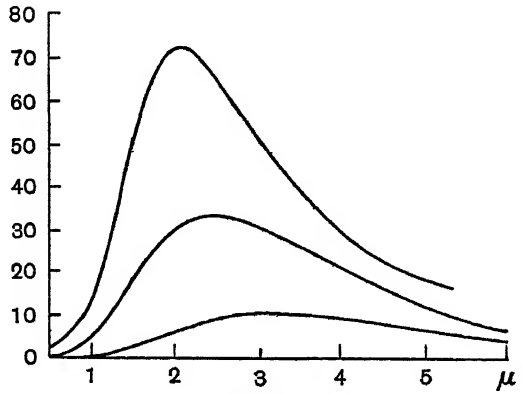


Fig. 88.

We shall see below that waves are now known of much longer and also of much shorter wave length than any of those hitherto considered.

Polarization

In sea-waves the particles oscillate in circular paths, in sound they move to and fro in the direction in which the wave is travelling. Hitherto in this chapter we have said nothing about the nature of the waves of light; the facts described are consistent with any wave theory. We must now deal with evidence that shows that the vibrations are transverse to the direction of propagation.

If light be passed through a plate of a crystal silicate called tourmaline, it acquires a slight colouration, though to the eye it otherwise looks unaltered. But, if a second plate of tour-

maline be interposed in the path of the light which has come through the first plate, differences appear. When the two plates are parallel, the light from the first passes freely through the second, but, if one plate be rotated round the beam of light as axis (Fig. 89) the light gets fainter, and, when the two plates are crossed at right angles no light gets through. Hence the light which passes through the first plate has acquired some new property, which prevents it from passing a plate at right angles to the first.

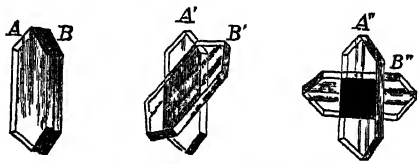


Fig. 89. Two plates of tourmaline.

If the vibrations which constitute light were longitudinal like those of sound, that is, if they were in the direction of propagation, no explanation of the effects with tourmaline could be offered. But, it is otherwise if they are supposed to be transverse. The transverse wave running along a rubber cord vibrated by hand, would traverse a slot cut in a plate if that slot were parallel to the direction in which the vibrations occur, but it could not pass a slot cut at right angles to the first. Hence the light which traverses the first plate of tourmaline may be analogous to the transverse wave running along a cord. If its vibrations, too, are transverse and take place in one plane, they might get through a crystal structure in which the vibrating particles can move in that plane, but not one in which any possibility of vibration is in a direction at right angles.

Ordinary light is supposed to vibrate transversely, the vibrations in the transverse plane taking place in all sorts of directions. The light which passes through a plate of tourmaline, on the other hand, vibrates transversely and all in one plane. It is said to be plane-polarized.

In tourmaline, only light polarized in one plane gets through, but, in other crystals, such as those of Iceland spar, a form of calcium carbonate, double refraction occurs and incident light is divided into two rays, one conforming to the ordinary laws of refraction and one not. These two rays are

found to be polarized in planes at right angles to each other. By cutting a suitable crystal of Iceland spar from A to A' as shown in Fig. 90, and cementing it together again with Canada balsam, we get what is called a Nicol's prism. If a beam of light SP fall on it, the ordinary ray O is totally reflected (see p. 203), and is turned aside, while the other, the extraordinary ray E, gets through. Owing to the colour absorption of tourmaline, a Nicol's prism is better for experiments on polarization.

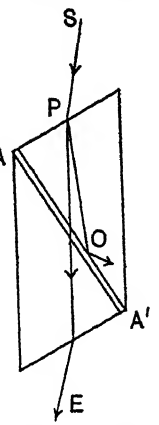


Fig. 90. Nicol's prism.

In the polariscope or polarimeter, two Nicol's prisms are mounted on the same axis with a space between them. If they be arranged with the crystals in the same position, light passes through and can be seen in a telescope placed beneath the second prism. If one prism be rotated through a right angle, the light is extinguished. Placing certain substances between the prisms is found to restore the light till the second prism has been turned through a further angle. Such substances are said to rotate the plane of polarization. This effect is of great theoretical interest in the history of chemical constitution (see p. 157) and of high practical importance in sugar analysis: it is the easiest way of measuring the amount of sugar in a given solution. When specially improved and adapted for this purpose, the instrument is known as a saccharimeter.

Electromagnetic Waves

In the section on electric units (p. 115) we saw that an electromagnetic wave, travelling through a medium of di-electric constant k and magnetic permeability μ, would move with a velocity $v = \sqrt{\frac{1}{k\mu}}$. By comparing the electrostatic and the electromagnetic units of some of the electric quantities, the value of this v has been determined. It is equal to 3×10^{10} cms. per second, a number equal to the observed velocity of light. Hence in 1864 Clerk Maxwell concluded that light was an electromagnetic wave of exceedingly short wave length.

The next step was clearly to demonstrate the existence of waves produced by electric means and to show their similarity to light. That electric and magnetic forces are exerted across space is clear from the fundamental experiments, and, if an electric current in one coil be continuously reversed in direction, it will induce another alternating current in a second coil at some distance from the first (see p. 107). Hence something analogous to wave alternations must pass across the intervening space.

The existence of definite regular trains of waves was first demonstrated experimentally in 1888 by Professor H. Hertz, then of Carlsruhe. A lightning flash or the discharge by sparking of a Leyden jar or other electric condenser is not a simple leakage of charge unless the resistance in circuit is very high. In general it consists of a series of electric oscillations which are gradually damped down by the resistance. The period of these oscillations may be guessed at by considering the analogy of the simple mechanical system of a weight hanging at the end of a spiral spring. As we saw on page 39, the period of this system is

$$T = 2\pi \sqrt{\frac{m}{F_1}},$$

where m is the mass of the vibrating system, and F_1 the force of restitution for unit displacement.

Starting or stopping an electric current in a coil of wire will, as Faraday first proved, induce a momentary electromotive force in a neighbouring coil, in the same direction as the primary current when that current stops and in the opposite direction when it starts. A similar E.M.F. will, of course, be induced in the primary coil itself, and it is clear that this induced E.M.F. will tend to oppose both the starting and the stopping of the current. Hence it is analogous to inertia or mass in mechanics. These inductive effects increase with the number of turns of wire in the coil and with the magnetic permeability of the substance of the core of the coil; they are proportional to what may be called the self-induction L of the coil, which is therefore analogous to m in the equation for the periodic time.

The quantity F_1 is the force exerted when the weight is

displaced by one centimetre against the elastic force of the spring. When an electric charge is given to a system, the potential produced, and therefore the opposing force, is greater the less is the capacity C of the system. Hence F_1 may be replaced by $1/C$, and it is likely that the period of these electric oscillations of the discharge from a condenser will be

$$T = 2\pi\sqrt{LC},$$

where C is the capacity of the condenser, and L the coefficient of self-induction of the circuit through which the discharge takes place. This guess is confirmed by both theory and experiment.

In a condenser made of two parallel plates near together, whether in the form of flat plates or of a Leyden jar, the lines of force mostly run straight from the units of positive charge on one plate to the negative units on the other; very few spread out into surrounding space. Hence such a system does not radiate much energy; the oscillations continue for a considerable time and only slowly die away. On the other hand, with two plates arranged as by Hertz (Fig. 91) or with simply two rods and sparking balls connected with the opposite terminals of an induction coil through choking coils, many of the lines of force run out into the surrounding space. Then, when a spark passes, the oscillations rapidly decrease in amplitude, because the energy is quickly radiated. The wave length of these oscillations is approximately twice the overall length of the rods and sparking balls.

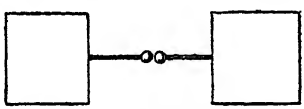


Fig. 91. Hertz's alternator.

The waves thus generated may be detected in various ways. Hertz used a resonator—a circle of wire with a gap at one point (Fig. 92). If this circle is made of such a size that the natural period of vibration of electricity on it is about the same as that of waves falling on the resonator at right angles to its plane, electric oscillations will be set up in the circle, and produce a visible spark between the balls on each side the gap. Other more practical types of receiver are described below.

Fig. 92. Hertz's resonator.

With a suitable electric oscillator and detector it is possible to show that electro-magnetic radiation behaves as would light of very long wave length. It can be reflected by "mirrors" made of large sheets of metal, and it can be refracted by prisms of pitch or other insulating material. If a grid made of parallel wires about half an inch apart is held between the transmitter and the receiver the waves get through unaltered when the wires are perpendicular to the oscillator rods, but are stopped if the grid wires are parallel to the oscillator rods, so that electric currents are set up in the wires. This experiment shows that these waves, like the light which has traversed a tourmaline crystal, is plane-polarized.

By placing a resonator between an oscillator and a reflector, Hertz showed that stationary undulations were produced. Just as a string attached to a tuning fork (Fig. 72, p. 190) is thrown into a series of nodes and loops, so the resonator will spark at certain points along the wave path and not at others. The wave length of the electric waves may thus be measured, and, if multiplied by the frequency, indicates a velocity equal to that of light.

A vivid picture of the propagation of an electromagnetic wave may be obtained by looking at the problem by the light of Faraday's conception of lines or tubes of electric force. As we have seen on page 94, the properties of magnetic fields of force can be represented by lines drawn everywhere in the direction of the force. In the same way, an electric field may be supposed due to electric lines or tubes of force in a state of tension. Each unit of charge, say each electron, may be imagined as the origin of one tube of electric force.

If an electron oscillates to and fro, a transverse tremor will run out along its tube, just as a wave runs out along a stretched cord when the end is moved by hand. The motion of a point on the cord is a transverse motion in one plane, and such a motion runs along the electric tube with the velocity of light, forming the wave-front. The electric force is everywhere in the direction of the tube, hence, if the wave-front be represented by the plane of the paper, that is, if the wave be travelling from the book towards the eye, the electric force will be in some line such as EE' in that plane, in the direction

OE when the vibration is in one phase and in the direction *OE'* when it is in the other. The motion of an electron constitutes an electric current, and a current causes a magnetic force at right angles to its direction. Hence the magnetic force in the advancing wave-front must lie in the line *MM'* at right angles to *EE'*, sometimes in the direction of *OM*, sometimes in that of *OM'*.

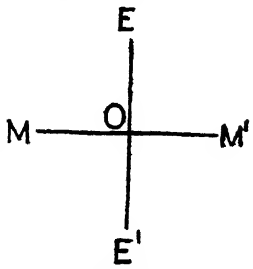


Fig. 93.

Electric tubes of force are perhaps only a convenient way of picturing what occurs in the electric field. On the other hand, an electron may be only a convenient figment of the imagination, standing in the mind as a symbol of the free end of an electric tube of force. However this may be, it is curious that, while Röntgen rays will ionize a gas, they affect only perhaps one molecule in a million. This suggests that these X-rays, which are but very short waves of light, do not give a continuous wave-front which would strike all the molecules, but consist of scattered units, somewhat, it may be, like tremors running along isolated tubes of force.

Wireless Telegraphy and Telephony

The electromagnetic waves we have been considering are, of course, now of great practical importance in telegraphy and telephony. The earliest successful attempts at wireless telegraphy were made with damped vibrations, for only thus could enough energy be radiated with the apparatus then available. The oscillator consisted of an aerial wire attached to one sparking ball and terminal of the secondary of an induction coil, the other ball and terminal being earthed (Fig. 94). With this apparatus, when a spark passes, an oscillation is set up. The energy is rapidly radiated, and the amplitude quickly diminishes in a manner represented in Figure 95. The next spark produces a similar oscillation of a few swings, and so on, each spark sending out a short train of waves like that shown in the figure.

Hertz's resonator has been already described, but it is not suitable for technical use. A more practical type of detector depends on the fact that the resistance at an imperfect

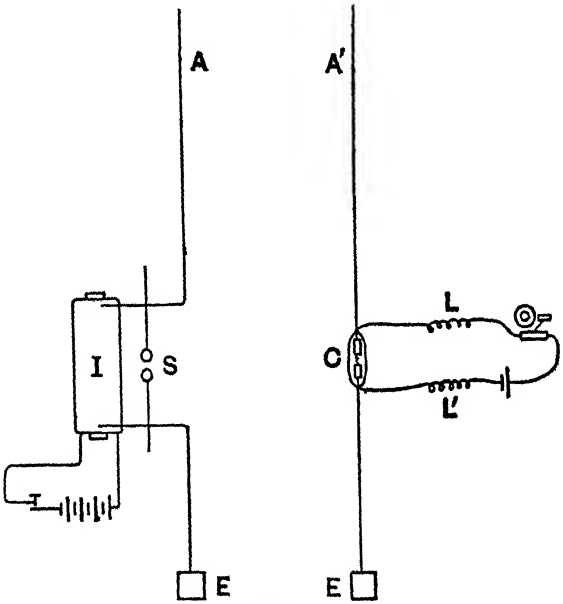


Fig. 94.

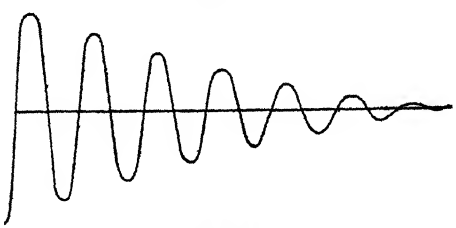


Fig. 95.

metallic contact is much diminished when electric oscillations pass across it. If metallic filings be placed between two poles in a glass tube, and the poles be connected with a battery and galvanometer, no current passes: the resistance of the

filings is too high. If, however, the poles are also connected, one with an aerial wire and one with earth, when electric waves arrive the oscillations have enough electric momentum to pass from filing to filing. Once a current is started, the filings cohere together, the resistance is greatly decreased, and a current will flow through the galvanometer till the coherence is destroyed by shaking the tube containing the filings. This apparatus is called a coherer.

The same principle is involved in the crystal receiving sets now used in wireless telephony. A flexible copper wire is pressed against a specially treated crystal of galena. The resistance at the point of contact is diminished by electric oscillations and much more in one direction than the other. This curious fact enables us to use a crystal contact as a rectifier—that is, an apparatus that only allows current to pass in one direction. Electric oscillations falling on it, therefore, are converted into intermittent gushes of electricity in one direction only, and will then deflect the needle of a galvanometer or, when intercepted at suitable intervals, produce sound in a telephone, neither of which can be done by the original oscillations.

Finally we have the instrument known as the thermionic valve. Certain incandescent solids, especially in a high vacuum, emit a stream of the negative corpuscles or electrons which have already been described as forming cathode rays. This effect is specially marked with the incandescent filaments of such substances as tungsten in electric lamps.

If such a filament, kept incandescent by a current, be connected with one terminal of a battery while the other is connected through a galvanometer with a nickel plate or cylinder inside the lamp, it is easy to show that a current passes from the filament to the plate when the filament is connected with the negative electrode of the battery and not when it is joined to the positive electrode. Negative electricity is carried easily by the projected electrons, but positive electricity cannot so escape. Hence this valve may be used as a rectifier.

The valve is improved by inserting another cylinder of metal gauze or coiled wire, called a grid, between the filament

and the surrounding nickel cylinder. When the grid is negatively electrified, the repulsion of similar charges prevents the electrons from the filament passing through to the outer plate. When the grid is positive, it helps the emission. The plate current increases as the positive potential of the grid is raised, till a certain limiting or saturation current is reached. If oscillations then fall on the valve, the swing of charge one way cannot increase the plate current which is already at a maximum, while the swing the other way does diminish it. Hence the effect of oscillations is always to decrease the plate current; this causes a telephone in the circuit to sound at the instant when the oscillations begin. If successive groups of oscillations, each group like that in Figure 95, fall on the valve at the rate of some hundred groups a second, the telephone will emit the corresponding musical note, and long and short signals, sent as recurrent groups of damped oscillations, may thus be detected and interpreted by the Morse code.

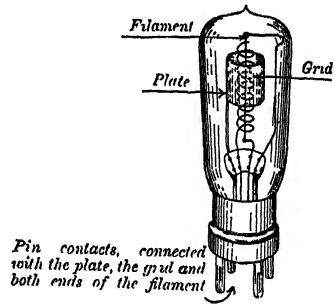


Fig. 96.

But in recent "wireless" practice damped oscillations are discarded in favour of continuous waves, in which the amplitude of vibration is maintained constant so long as the waves last. It is possible to do this with a primary oscillating system in which oscillations can be maintained undiminished, and to use these oscillations to induce secondary oscillations in another system which is a good radiator. Large quantities of energy can thus usefully be employed, and continuous wave radiation be maintained. Damped oscillations will affect almost any detector, since the first swing is so much larger than those few that follow, but with continuous oscillations a receiver can be constructed of definite period of vibration which, on the principle of resonance, will respond only to waves of the same definite period. By thus "tuning" transmitter and receiver, waves of definite wave length, that is, of

definite period, can be used for one purpose, independently of other waves traversing the same space.

The thermionic valve described above can be used not only as a detector, but also as a source of continuous undamped electric oscillations. The primary coil of an induction coil is inserted in the plate circuit of the valve, and the secondary coil coupled with the grid and the filament. This arrangement is found to start self-oscillation, the alternating potential of the secondary circuit varying the plate current and thus maintaining the alternations of potential.

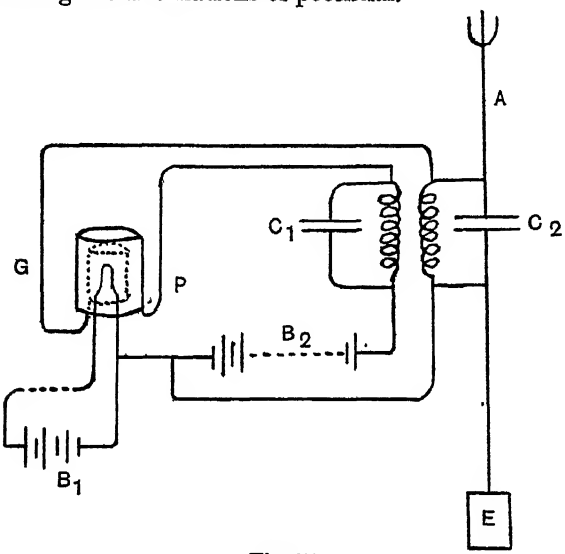


Fig. 97.

A diagram of one arrangement of connexions is given in Figure 97. The incandescence of a filament is maintained by the battery B_1. The filament is also joined to the negative electrode of the high tension battery B_2, the positive electrode being connected with the plate P through the primary coil and condenser C_1. The negative electrode of the plate battery is also joined with the grid through the secondary coil and a second condenser C_2. The frequency of the oscillations can be adjusted by altering the capacities C_1 and C_2.

This arrangement gives a simple harmonic electric oscillation, as a tuning fork gives a simple harmonic sound wave. A number of valves may be coupled together to increase the energy of the oscillations.

The oscillations can be used to emit continuous waves by connecting one plate of the condenser C_2 with earth, and the other with one or more aerial wires which are generally carried some distance vertically and then horizontally. This aerial system forms a condenser with the earth, and the capacity of this condenser is added to that of C_2.

By placing the coils of transformers in the grid and plate circuits, a thermionic valve may be used as an amplifier; a weak electromotive force, due to feeble oscillations in an aerial wire, is found to be thus much intensified. The success of long distance wireless telegraphy largely depends on this amplifying effect.

The general principle of wireless telephony depends on emitting a regular train of continuous waves called carrier waves, modifying this train intermittently by the action of a sound at one end, and reproducing sound at the other. Thus, in a sense, it is the carrier wave that acts as the wire acts in ordinary telephony; the impulses that reproduce the sound run along the wire in one case and the wave in the other. The wave itself while unaltered does not affect the telephone, the vibrations are much too rapid—say 750,000 a second, instead of the 100 to 3000 which is about the range of audible sound. But, if recurrent changes are impressed on the wave, at any rate between the audible limits, those changes will affect the telephone.

An apparatus like that illustrated in Figure 97 may be used. A loud speaking telephone is coupled inductively through an induction coil and another valve circuit with the grid circuit. The vibrations of speech in the microphone are thus translated into corresponding periodic variations of potential in the grid, and this results in variations in the continuous waves emitted by the aerial wires, the variations being of similar wave-form to that of the sound.

At the receiving station, the waves, much weakened by distance, are caught by another aerial, and the variations of

the carrier waves are reconverted into speech, the circuits being timed by adjusting the induction and the capacity so as to give vibrations in unison with the waves to be received. The standard wave lengths of carrier waves for "broadcasting" lie between 350 and 425 metres. The velocity being 300,000 kilometres per second, the frequencies lie between 857,000 and 706,000 alternations per second. On these waves variations are impressed at "audio-frequencies" of perhaps 100 to 1000 per second, and reproduced as sound in the receiving telephones.

Röntgen Rays or X-rays

The wave lengths used in wireless telegraphy are far longer than those of visible light. We must now describe another set of waves which are much shorter than those of light.

In 1895 Professor Röntgen of Munich, working with electric discharges in vacuum tubes, noticed that covered photographic plates became fogged if kept in the neighbourhood of the tubes. This accidental discovery may be said to have been the starting point of modern physics.

Whenever cathode rays impinge on solid objects, the impact of the electrons starts a stream of X-rays. This type of radiation, when produced in high vacua, is absorbed by matter in proportion to the density. Hence, the rays traverse skin and flesh, and are stopped by metal and bone, to the great gain of modern surgery. They cause fluorescent screens to glow, and can therefore be made visible to the eye in a dark room. They affect photographic plates, and their incidence can thus be permanently recorded.

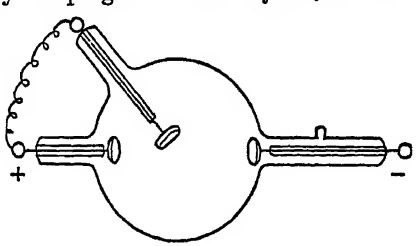


Fig. 98.

A general form of X-ray bulb is shown in Figure 98. The cathode, or negative electrode, is a concave aluminium cup which focusses the cathode rays on to a second aluminium plate or anti-cathode placed near the middle of the bulb. Thus the anti-cathode is bombarded by a stream of cathode rays

and becomes the source of X-radiation. The anode may be placed anywhere in the bulb. The higher the exhaustion of air, the more penetrating or "hard" the X-rays become.

Röntgen rays are not refracted like ordinary light, and very little trace of regular reflexion or of polarization can be detected. But, on the other hand, they are not deflected by magnetic or electric forces like cathode rays. Their nature was for some time a subject of discussion, but finally Laue made the remarkable discovery that the atomic structure of crystals will give X-ray spectra as finely ruled gratings will give a spectrum with visible light. This new field of research enabled Sir William Bragg and his son, W. L. Bragg, to prove that the periodic structure in crystals is due to successive layers of atoms. They measured the dimensions and form of the atomic arrangement in various crystals, and showed that the wave length of the X-rays was of the order of 10^{-8} cm.

We may now collect together the information available about the ranges of the wave lengths of electromagnetic radiation.

Range	Radiation	Wave lengths in centimetres
?	X-rays	10^{-8}
4 octaves	Ultra-violet light which can be photographed	$0{\cdot}25 \times 10^{-5}$ to 4×10^{-5}
1 octave	Visible light	$4{\cdot}0 \times 10^{-5}$ to $7{\cdot}6 \times 10^{-5}$
8 octaves	Infra-red radiant heat	8×10^{-5} to 3×10^{-2}
8 octaves	Hertzian waves	5 to 1000
11 octaves	Wireless telegraphy & telephony	100 to 2 million

The diffraction spectra of X-rays produced by crystals used as gratings consist of a mixture of diffuse radiation of all wave lengths within certain limits, and of more intense radiation of definite frequency superimposed as spectral "lines" upon it. This characteristic line radiation is a diffraction phenomenon similar to the line spectra obtained with visible light, and with it a very important discovery was made in 1913 and 1914 by a young Oxford man, H. G. J. Moseley, who, soon afterwards, was killed in the war.

When the target bombarded by cathode rays was changed from one metal to another, and the spectrum of the resultant X-rays examined by using a crystal of potassium ferrocyanide

as a grating, Moseley found that the frequency of vibration of the characteristic lines in the spectrum undergoes a simple change. The square root of the number n, expressing the number of vibrations per second corresponding to the strongest line in the X-ray spectrum, increases by the same amount on passing from element to element in the periodic table. If $n^{\frac{1}{2}}$ be multiplied by a constant, so as to bring this regular increase to unity, we get the series of atomic numbers, already given in chapter I, page 22. They range regularly for all solid elements examined, from aluminium 13 to gold 79. Filling in the other known elements, we find that, from hydrogen 1 to uranium 92, there are only five places left for undiscovered elements.

It appears then that only five ordinary elements with atomic numbers between 1 and 92 remain unknown. But, as we shall see presently, radioactive elements have been found, identical with certain ordinary elements in atomic number but differing from them in atomic weight. Possibly more of these radioactive isotopes may yet come to light.

The accepted explanation of Moseley's most remarkable discovery of atomic numbers will be given in the next chapter.

CHAPTER VIII

RADIOACTIVITY, ATOMIC RADIATION AND RELATIVITY

The Discovery of Radioactivity

RÖNTGEN rays produce light when they strike screens powdered with fluorescent substances like zinc sulphide or barium platino-cyanide. Hence it was natural to enquire whether conversely any fluorescent substances might themselves give forth X-rays or other similar rays spontaneously. Led by this idea, Becquerel discovered in 1896 that compounds of uranium, whether luminous or not, emitted rays which would affect photographic plates and fluorescent screens, and it was soon found that thorium and its compounds did likewise.

These newly discovered rays cause electric conductivity in air or other gases through which they pass, in this resembling X-rays, and this fact gave another and much more delicate test for radioactive bodies.

In the year 1900, M. and Mme Curie made a systematic search for these effects in a great number of substances, both natural and artificial. They found that pitchblende, a mineral containing uranium, was more active than uranium itself. This suggested the presence of some more active constituent. The components of pitchblende were therefore separated from each other by chemical analysis, and three new radioactive elements were thus discovered and named radium, polonium and actinium.

Radium, the most remarkable of these new elements, is first obtained in company with barium, which it resembles chemically. But there is a slight difference in the solubility of some of their salts, and hence they may be separated, but only after long and tedious processes of fractional crystallization, tons of the mineral yielding only a small part of a gram of an impure salt of radium.

As the introduction of the new methods of electrolysis and

of spectrum analysis led to the discovery of new elements in 1807 and 1860 respectively, so radioactivity has again extended our knowledge of the varieties of matter. The leak of electric charge from an electroscope, due to the passage of radioactive rays through the surrounding air, is an extremely delicate test. It will show the presence of less than the thousandth part of the substance needed to give a visible spectrum by the most refined methods of spectrum analysis.

The series of discoveries of new elements by new methods in the past naturally makes us prone to look for an indefinite repetition of such successes, but Moseley's proof of the existence of atomic numbers, with nearly all places already filled, shows that, between hydrogen and uranium, at all events, the limits of discovery are nearly reached, except perhaps as regards new radioactive isotopes with the same atomic numbers as known elements but different atomic weights.

The Theory of Radioactivity.

In 1899 Sir Ernest Rutherford, then Professor Rutherford of Montreal, began those epoch-making researches, on which, combined with Sir J. J. Thomson's work on cathode rays, modern physics is founded. To Rutherford and his pupils we chiefly owe our knowledge of the processes and meaning of radioactivity, knowledge which we will now shortly summarize.

Radioactive substances emit more than one kind of radiation:

(1) α-rays produce marked effects on the electric conductivity of gases; they are slightly deflected by magnetic and electric fields, and have hence been shown to be particles of the mass of helium atoms, each carrying a positive charge of two units;

(2) β-rays affect photographic plates through opaque screens; their magnetic and electric deflections are large, and prove them to be similar to cathode ray particles of unusually great velocity;

(3) γ-rays are still more penetrating; they are not deflected, and resemble Röntgen rays of the characteristic "hard" type obtained in very high vacua.

Wherever there is radioactivity, it is found that new chemical substances are being formed. If α-rays are allowed to accumulate, the spectrum of helium soon appears. This confirms the evidence given by their magnetic and electric deflexions—α-rays are projected helium atoms carrying positive charges.

Uranium is found to be continually developing minute quantities of a different body called uranium X, which may

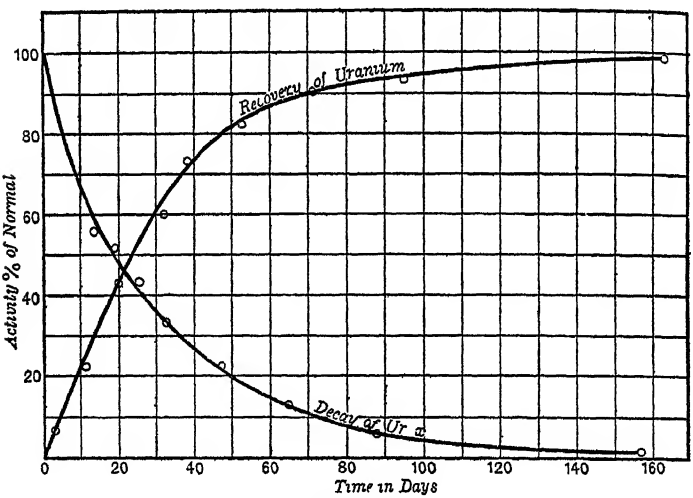


Fig. 99.

be separated chemically. This body is very active photographically, while the uranium from which it has been separated, though it still ionizes gases, is for the time photographically inert. But, if kept, the uranium X slowly loses its activity, and the original uranium regains it.

A similar phenomenon is found with radium, which produces a gaseous emanation. Most of the gas is stored in the solid radium salt and can be driven off by heat. It is radioactive, and the parent radium loses an equivalent amount of radioactivity when the emanation is removed.

The rate at which the activity I of uranium X or radium

emanation decays is proportional to the amount of activity remaining, that is

$$-\frac{dI}{dt} = \lambda I.$$

Hence we see that the law is that of the mono-molecular reaction (see p. 147). It is illustrated in Figure 99, which represents the decay of the activity of uranium X and shows also the recovery of activity in uranium. The activity of uranium X decays to half value in 22 days, and that of radium emanation in 3·7 days. Other similar changes require times measured in seconds, or in thousands or millions of years.

In 1903 Curie and Laborde made the surprising discovery that radium compounds continuously and spontaneously emit heat. From experiments on the small quantities available it has been calculated that one gram of pure radium would emit 135 gram-calories per hour. Other evidence shows that the activity of radium only falls to half-value in about 1600 years. Hence it must continue to emit heat with little diminution for hundreds of years, and the total thermal output of one gram of radium would greatly transcend that associated with any ordinary chemical action.

Finally, the radioactivity of any substance is proportional to the amount of the radioactive element which it contains. One gram of uranium produces the same effect whether it be in the form of the metal or combined with other elements to form an oxide or a salt.

To sum up, we may say: (1) that the chemical change associated with radioactivity is mono-molecular, the decomposition of single particles and not a combination; (2) that the activity is proportional to the mass of the radioactive element whether free or combined, so that it is an affair of atoms and not molecules; and (3) that the amount of energy liberated is many thousand times more than that associated with the most violent known chemical reaction.

As a result of all this evidence, Rutherford and Soddy in 1903 framed their revolutionary theory that radioactivity is due to an explosive disintegration of the elementary atoms, and no other known hypothesis can explain the facts. Here and there one atom out of many millions suddenly explodes;

an α-particle, or a β-particle and a γ-ray are ejected, and a different atom is left behind. If an α-particle has been thrown out, this new element will be less in atomic weight by the four units of a helium atom.

As an illustration we may well trace the pedigree of the radium family. The earliest known ancestor is uranium, a heavy element of atomic number 92 and an atomic weight of 238. Its descendants are shown in the following table:

	Atomic number	Atomic weight	Time of half decay	Radio-activity
Uranium I	92	238	5×10^9 years	α
↓				
Uranium X_1	90	234	23·5 days	β, γ
↓				
Uranium X_2	91	234	1·17 minutes	β γ
↓				
Uranium II	92	234	2×10^6 years (?)	α
↓				
Ionium	90	230	2×10^5 years (?)	α
↓				
Radium	88	226	1590 years	α
↓				
Radium Emanation...	86	222	3·85 days	α
↓				
Radium A	84	218	3·05 minutes	α
↓				
Radium B	82	214	26·8 minutes	β, γ
↓				
Radium C	83	214	19·5 minutes	α, β, γ
↓				
Radium D	82	210	16·5 years	β, γ
↓				
Radium E	83	210	4·85 days	β, γ
↓				
Radium F (Polonium)	84	210	136 days	α
↓				
Lead	—	206	Inactive	

Uranium emits an α-particle, a helium atom of mass 4 and charge $+2$. Uranium X_1 remains behind: its weight is $238 - 4 = 234$, and its atomic number $92 - 2 = 90$. Its radioactivity gives β- and γ-rays only; the β-particle being of negligible mass and carrying a unit negative charge. Hence the body into which it passes, known as uranium X_2, has one negative charge less, that is, one positive charge more, than Uranium X_1 and consequently an atomic number of 91. Its

atomic weight is unaltered, 234. Its radioactivity also consists only of β- and γ-rays; hence its child, uranium II, has an atomic number of 92 and, again, the same weight of 234.

And so we go on as shown in the table. When an α-ray is emitted, the product has an atomic weight less by 4 units and an atomic number less by 2. When a β-ray comes off, the weight is unchanged, but the number is raised by unity.

The last known descendant of the family is lead. Lead, as obtained from radioactive minerals, has been shown by Soddy to be appreciably different in atomic weight from ordinary lead, though identical with it in chemical properties. This is to be compared with Aston's work on positive ray mass-spectra, described on page 122. The general conclusions are (1) that the chemical properties of the elements depend on their atomic numbers; (2) that pairs of elements called isotopes exist, which have the same atomic number, and therefore the same chemical properties, but different atomic weights; and (3) that many common elements consist of mixtures of these isotopes.

The Individual Atom

The atomic theory was established by the chemical work of Dalton, but for a hundred years it was impossible to point to any demonstration of the existence of single atoms; they could be treated only statistically by millions. But radioactivity now makes it possible to trace the effect of single α-particles. Their effect was first seen by Crookes as scintillations on a fluorescent screen of zinc sulphide, and other methods of detection are now available.

A gas at a pressure of a few millimetres of mercury, when exposed to an electric field just weaker than that needed to cause a spark, is in a very sensitive state. An α-particle, owing to its immense velocity, will produce many thousand ions by collision with the molecules of gas. These ions are subject to strong electric forces, are therefore set in rapid motion and thus produce other ions by collision. Hence the total effect of a single α-particle is multiplied, and the needle of a sensitive electrometer may be made to give a throw of 20 millimetres or more on the scale. By using a very thin film of active

matter, Rutherford reduced the throws to three or four a minute, and counted the number of α-particles projected. The life of radium may thus be estimated. A mass of radium diminishes by one half in about 1600 years.

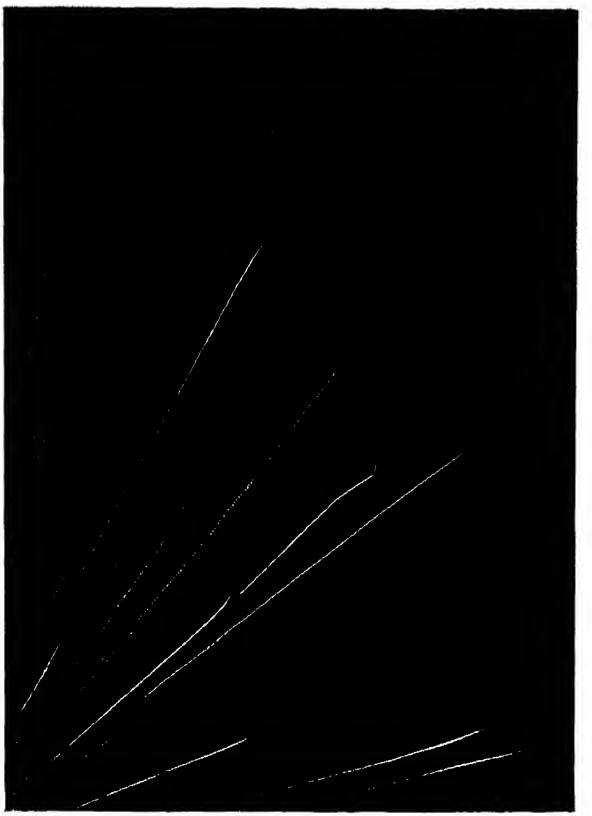
Another method is due to Mr C. T. R. Wilson. When shot through moist air saturated with water-vapour, α-particles produce ions which act as nuclei of condensation. Cloud-tracks are thus formed, showing the path of each α-particle. They may be photographed as shown on Plate III.

The Structure of the Atom

Sir J. J. Thomson's fundamental discovery of the cathode ray corpuscle, the same in all kinds of matter, suggested that this corpuscle was a common constituent of all chemical elements. Since light had been proved to be an electromagnetic wave, the emission of definite line spectra by many elements indicated that in them electric units must exist, capable by their vibration of starting the undulations of light. The short periods of these undulations show that the vibrators must be much smaller than atoms, and thus an electric theory of matter had been put forward before Thomson's discovery by Lorentz and Larmor, who used Stoney's name of electron for the electric vibrating units. These electrons may clearly be identified with Thomson's corpuscles.

The corpuscles carry negative charges and are therefore negative electrons, and, if they form part of the structure of neutral atoms, an equivalent positive charge must also be present therein.

The cloud-tracks of α-particles are usually straight, but an occasional sharp change in direction occurs: one such deflexion is shown on Plate III. The forces exerted by negative electrons on an α-particle must be too small to produce such a change, but the effect is explained if it be supposed that an atom is a complex body of very open structure having a positive charge concentrated in a minute nucleus, with negative electrons revolving round it in planetary orbits. Since a normal atom is electrically neutral, the positive charge in the nucleus must be equal as well as opposite to that on all the electrons together, and, since the mass of the electrons



CLOUD-TRACKS OF α PARTICLES

is small compared with that of an atom, nearly all the mass must be concentrated in the nucleus.

A moving electric charge carries a field of electromagnetic force with it, and this, since it has energy, must possess inertia. Hence an electric charge has something which behaves like mass, which, indeed, cannot be distinguished from mass, and may be the real meaning of matter. Mathematical analysis shows that, unless the particle is moving with very great speed, the electric mass is $2e^2/3r$, where e is the charge and r the radius. Hence by making r small, that is, by concentrating the charge, the effective mass is increased. The mass of the positive nucleus of hydrogen, that is, of the atom, being 1700 times the mass of a negative electron, the radius of the positive nucleus will be only the 1700th part of the radius of an electron. This positive unit, which corresponds to the electron as the negative unit, is sometimes called a "proton."

A hydrogen atom consists of a unit positive nucleus or proton, with a single negative electron revolving round it in an elliptical orbit, like a planet round the sun. Other atoms consist of a complex nucleus containing N positive units, where N is Moseley's atomic number, bound together by a certain number of negative electrons, and having other electrons moving round this centre. N also represents the total number of electrons in the atom, and Ne the total negative charge on those electrons, neutralized by an equal positive charge in the nucleus. As α-rays are flights of helium atoms, it is probable that helium atoms are some of the bricks with which the nuclei of other heavier atoms are built up. The helium atoms themselves must each be made up of four protons or hydrogen nuclei, too firmly joined to be separated even in the adventurous life of an α-particle.

Since atoms can be ionized, and given one, two, three or possibly four unit charges in accordance with their chemical valency, a small number of electrons can be added to or subtracted from an atom with no fundamental change in its nature. We may suppose these electrons to move in outer orbits, while others perhaps revolve in inner rings, and yet others form an essential and, in general, a stable part of the nucleus.

As we have explained, most radioactive transformations are associated with the ejection of an α-particle, which is a helium atom of mass four, carrying two units of positive electricity. They must therefore involve a catastrophic change in the nucleus. The residue will be four units lighter, and two negative electrons must be discarded to re-establish neutrality: a new atom, a new element, will result.

Thus the transmutation of matter, the dream of the medieval alchemist, has at length been demonstrated. Till lately, no human means had been discovered of hastening, still less of controlling, these changes. They depend on chance happenings within the atom, and their frequency conforms to the well-known laws of probability. But Rutherford has now discovered that a bombardment by α-rays will induce radioactive transformations in certain atoms such as those of nitrogen. Nitrogen has an atomic weight of 14, made up of three helium nuclei weighing 12, and two odd hydrogen nuclei. When struck by an α-particle, the nitrogen nucleus is shattered, and, of its constituents, hydrogen nuclei are shot out with great velocity. Here, then, we catch a glimpse of the possibility of inducing at will a breakdown of the atom—transmutation in one direction. But it is easier to destroy than to build up: it does not follow that we shall ever be able to construct the heavier and more complex atoms from the lighter and simpler.

Atomic Radiation

If an electron oscillates backward and forward in a line, it emits electromagnetic waves, and if a number of electrons so oscillate together, as when an alternating current oscillates in an aerial wire, the waves are emitted in all directions at right angles to the line, so that the wave-front is an ever-growing cylinder with the line as its central axis. If an electron describes a circular planetary orbit about an atomic nucleus, it again has acceleration towards a centre, its direction of motion changing, though not its velocity. This acceleration also would naturally involve the radiation of energy, and, on Newtonian dynamics, the effect would be a contraction of the orbit, with a consequent quickening of the period of rotation and of the frequency of the waves

emitted. Atoms at all stages of this process would exist, and therefore in all spectra radiation of every frequency should be found, instead of one or a few definite and unchanging frequencies only, characteristic of the line spectra of the atoms of many chemical elements.

Again, the energy in the spectrum of an incandescent solid is not evenly distributed, but, as illustrated in Figure 88 on page 211, is greatest for certain frequencies in the invisible infra-red radiation. It is difficult to see how this could be, on any ordinary and obvious theory of atomic or electronic radiation. Indeed, mathematical investigation indicates that oscillators of high frequency should radiate more energy than those of low, so that visible light should give more heat than the invisible infra-red rays, and ultra-violet rays more than light. All this is contrary to well-known fact.

To meet these difficulties, Planck devised a "Quantum Theory," according to which radiation is not continuous, but, like matter, exists only in indivisible units or atoms. The units are not all the same size, but are proportional to the frequency of the oscillation. Hence high frequency ultra-violet oscillations can only radiate when they have a large amount of energy available, and therefore the chance of many units being available and radiated is very small, as is the total energy emitted likewise. The low-frequency oscillations, on the other hand, emit small units, and therefore the probability is great that many units will be available and radiated; but, since each unit is very small, the total amount of energy is small also. For some special range of intermediate frequency, where the chances are favourable, the unit is of medium size, the number radiated may be fairly large and the total energy a maximum. Observation shows that, for temperatures available in terrestrial laboratories, this optimum range for radiation is below the visible spectrum in the higher part of the infra-red region.

It is a problem in probability analogous to that of raising money for a testimonial by a subscription of uniform amount. If the subscription be made too high, say ten pounds, the chances are that very few people will join, and the total amount raised will be small. On the other hand, if the sub-

scription were fixed at one shilling, the probability is that many shillings will come in, but, since a shilling is a small sum, the total amount raised is again small. But, with a certain intermediate unit dependent on the nature of the field and the popularity of the proposed recipient, let us say with the unit fixed at a modest guinea, a considerable number will be found willing to subscribe or anxious not to be left out, and the total sum raised will be a maximum.

To explain the facts, Planck's quantum of energy ϵ must be supposed proportional to the frequency or inversely proportional to the period of vibration. Hence we see that

$$\epsilon = hn = \frac{h}{T},$$

where n is the frequency, T the period and h a constant. Hence Planck's constant h is $T\epsilon$, the product of energy and time, a quantity which is called Action. This constant unit of action is, of course, independent of the frequency, indeed of everything variable. It is a true natural unit, analogous to the natural unit of matter and electricity found in the electron.

A theory made specially to suit one definite set of facts can be adjusted to fit those facts, and, however good the fit and fashionable the cut, the evidence for the theory being universally useful is not perhaps very high. But, if an entirely different set of phenomena can be explained by the same theory, especially if no rational account of them has been given otherwise, the value of the evidence is very greatly increased, and we begin to feel confident that the theory may be relied on to co-ordinate still more relations.

Planck's theory was introduced to meet the facts of radiation. As it involved a breach with orthodox dynamics, it was rightly viewed with caution, if not scepticism. But when it was applied successfully by Einstein to explain the phenomena of specific heat, the probability of its extended usefulness was very much increased.

The ordinary kinetic theory (see p. 74) suggests that monatomic molecules in solids should have an atomic heat of three times the gas constant, or nearly six calories per

degree, and that this quantity should be independent of temperature. Metals contain monatomic molecules; at ordinary temperatures their atomic heats are roughly constant and equal to 6. But at lower temperatures the value diminishes.

This was first successfully explained by Einstein, who pointed out that, if energy could only be absorbed by definite units or quanta, the rate of absorption would depend on the size of the unit and therefore on the frequency of vibration and thus on the temperature. He calculated a formula, which gave results agreeing with observation—markedly so in the case of carbon, which, even at ordinary temperatures, has a variable atomic heat, much below the value for metals.

On this quantum theory, light is neither the steady aethereal wave of Fresnel nor the continuous electromagnetic wave of Maxwell, but is emitted in minute gushes of electromagnetic energy which may almost be regarded as atoms of light, equivalent to, though different in kind from, Newton's corpuscles. But the reconciliation of this conception with the phenomena of interference remains a difficulty for the future.

The application of Planck's quantum theory to the problem of atomic structure was made by Niels Bohr, now of Copenhagen, a brilliant young physicist who had worked with Rutherford. To explain the general stability of atoms and the existence of spectra with sharp lines, Bohr made two postulates:

(1) That among possible states of motion in an atom there are a number of states possessing a stability which is not explicable by ordinary mechanics, and that any permanent change in the motion must involve a complete transition from one stable state to another;

(2) That, contrary to ordinary electromagnetic theory, no radiation is emitted by the atom in its stable states, but that, when an abrupt change occurs from one stable state to another, a sudden and limited radiation is emitted such that the change in energy in the atom $E_2 - E_1 = hn$, where h is Planck's constant, and n the frequency. Conversely, an absorption of energy may reverse the process and carry the atom back to its former state of stability.

On these assumptions, Bohr imagined a model of the hydrogen atom with a positive nucleus and one electron which revolves about the nucleus in one of four stable elliptical paths. He proved that this model would explain the known series of lines in the hydrogen spectrum, radiation taking place as the electron passed suddenly from one stable orbit to one or other of the rest, the frequency of the radiation being proportional to the amount of energy emitted.

This structure of the atom is illustrated in Figure 100. The concentric circles represent the four stable orbits, and the radii the six possible jumps from one of the orbits to another.

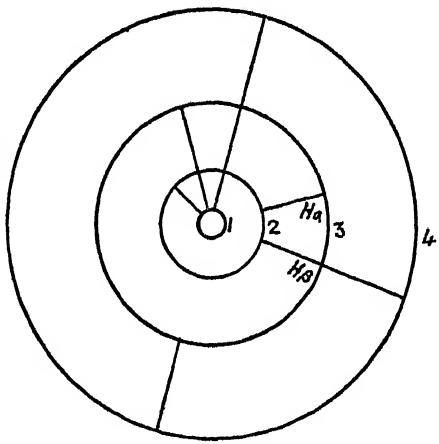


Fig. 100.

The Luminiferous Aether

The success of the wave theory of light indicated that space must be supposed filled with a medium to carry the waves, to supply, as the late Lord Salisbury once said, "a nominative case to the verb to undulate."

The polarization of light shows that the waves must be transverse; hence, if the waves are imagined to be waves of mechanical motion in a medium, that medium must possess rigidity, and this reasoning led for a time to elastic solid theories, in which a highly tenuous aether was given a rigidity far beyond that of steel.

Maxwell's electromagnetic theory and its experimental verification by Hertz led to a new outlook. The waves of light became electromagnetic in their nature. If a mechanical theory of a universal aether was to be maintained, electromagnetics must be expressed in terms of mechanical move-

ments or strains in that medium. On the other hand, it was equally possible to abandon the attempt to apply mechanical conceptions to the aether, and rest content with the words "electromagnetic waves" as a proximate explanation of light.

That this is unsatisfactory to our minds may perhaps be traced to the fact that we possess no special electric or magnetic sense. Possibly to the electric fish called the torpedo, which has a special organ for administering electric shocks, electric force may be a direct sense-perception, familiar and fundamental as mechanical force to ourselves. To a torpedo, the expression "an electromagnetic wave" may have a clear and simple and definite meaning.

And possibly we may find it useful to regard space as filled with a vast network of Faraday tubes of electric force, whatever they may be, instead of with a uniform elastic solid aether, or even to refrain from making pictures at all. We now know too much to dogmatize.

The Velocity of Light

The discovery that light needed time for its propagation was made by the Danish astronomer Olaus Römer in 1676. Römer found that the intervals between the successive eclipses of one of the satellites of the planet Jupiter were longer when the earth was receding from Jupiter and shorter when the earth was approaching. He estimated the velocity of light as 192,000 miles a second.

Fifty years later, James Bradley, the Astronomer Royal, got a concordant result from the aberration of the light from the fixed stars. As seen from a distant star in the plane of the earth's orbit, the earth would seem to oscillate from side to side once a year, moving in opposite directions in successive six months. The rays shot from the star to hit the earth must always be aimed a little in front of it as we shoot in front of a driven partridge or a rocketting pheasant, and so, if the star now shoots to the right of the earth's true position, six months on it must aim to the left. This means that the rays by which the star is seen from the earth at different times are not parallel to each other, but that the star appears to move periodically in space as the year revolves. From this

apparent movement, the ratio of the velocity of light to the velocity of the earth in its orbit may obviously be calculated.

The first determination of the velocity of light over short distances on the earth was made in 1849 by Fizeau, who passed a beam of light through one of the gaps in a toothed wheel, and reflected it back on its path by a mirror three or four miles away. When the wheel was at rest, the return beam passed back through the same gap and was visible on the other side, but, as the wheel rapidly revolved, a speed could be found at which the return way was blocked by the next tooth. The time occupied by the wheel in spinning through this small part of a revolution is clearly the time required for light to travel to the distant mirror and back again.

A better method is that of L. Foucault. A beam of light from a slit *S* is made very slightly convergent and then reflected from the plane mirror *R* to a focus on a concave mirror *M*. It returns along its path and forms an image of the slit on the slit itself. The mirror *R* is then rotated rapidly at a known speed. It moves through a small angle while the light travels from *R* to *M* and back again, and therefore the return path *RS′* is not coincident with *RS*, but will be turned through an angle twice the rotation of the mirror *R*. A plate of unsilvered glass is placed at 45° with the axis between *S* and the mirror, and this by reflected light forms two images of the slit, the distance between which can be measured.

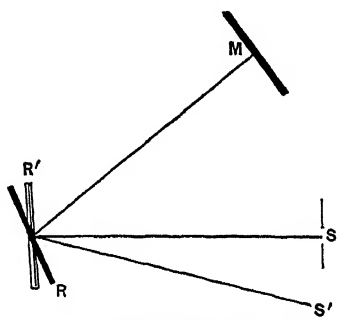


Fig. 101. Foucault's apparatus.

The best modern results for the velocity of light give a value of 186,326 miles or $2{\cdot}9986 \times 10^{10}$ cms. a second *in vacuo*, that is 3×10^{10} cms. per sec. within less than one part in a thousand.

If there is anything of the nature of a luminiferous aether, it seems natural that the effect on light travelling through it should make possible the determination of its motion. If the

earth moves through the aether without disturbing it, the earth and the aether will be in relative motion. In that case, light should be found to travel faster when it moves with the aether than against it, and on the whole faster when it travels to and fro across the aether stream than when it passes first in its direction and then against it. It is quicker to swim from one side of a river and back, than an equal distance up and down stream.

That is the essence of the famous experiment made by Michelson and Morley in 1887. They mounted their apparatus on a stone floating in mercury to prevent vibration. A beam of light *SA* is partly reflected and partly transmitted at the glass *A*. The two parts are reflected by mirrors at *B* and *D*. If $AB = AD$, the paths are equal in length, and interference effects will be seen in a telescope by an eye at *E*. Let us imagine that the earth is moving in the direction *SAD* but not carrying the aether with it, so that the aether is moving through the laboratory as the wind through a grove of trees. This will introduce a difference in the time of transmission over the paths *ABA* and *ADA*, and the interference fringes will not occupy the same place as they would if the aether and the earth were relatively at rest. Next let the apparatus be floated round through a right angle. *AB* is now in the direction of motion and *AD* across it. The interference fringes should now move in the opposite direction, the whole displacement being twice that noted above.

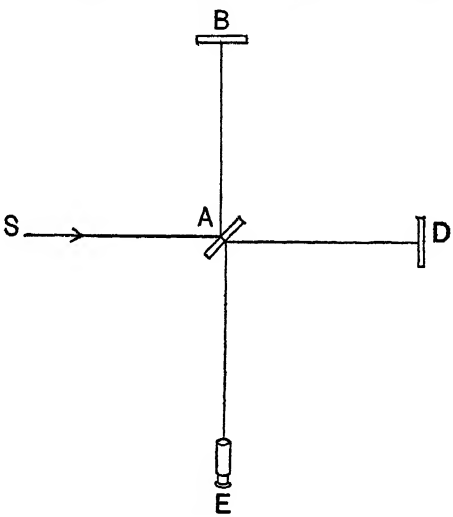


Fig. 102.

But Michelson and Morley could observe no measurable displacement of the interference fringes, and concluded that there is no appreciable relative motion of the earth and the aether. A repetition of their experiment showed that on their assumptions the relative motion is certainly less than the tenth part of the earth's velocity in its orbit. The earth seems to drag the aether with it.

In calculating the velocity of light from the observed aberration of stars, it is assumed that the aether is undisturbed by the motion of the earth through it. Moreover, Lodge in 1893 could find no change in the velocity of light between two heavy steel plates spinning at the highest safe speed. Hence masses of this size do not drag the neighbouring aether with them. Thus both the theory of aberration and the deductions from Lodge's experiment seem quite inconsistent with Michelson and Morley's result.

Whenever we get a discrepancy of this kind, if we are to hold to our belief in the uniformity and rationality of nature, we may conclude that something must be wrong either in our experiments or in our conceptions of the causes at work, and it is probable that an interesting and necessary revolution in ideas is under our eyes if we can but see it.

The first useful suggestion was made by Fitzgerald and developed by Larmor and Lorentz. If matter is electrical in essence, or, indeed, if it is bound together by electric forces, it may contract in the direction of the motion as it moves through an electromagnetic aether. Such a contraction would not otherwise be observed, firstly because it would be too small, and secondly because any scales we used to measure it would themselves be subject to the same contraction, so that in the direction of motion, the unit of length would be shorter. Thus Michelson and Morley's apparatus might change in dimensions as it rotated in such a manner as to compensate for the displacement of the interference fringes expected to be produced by the earth's movement through the aether.

It is easy to calculate the contraction necessary. Let c be the constant velocity of light through the aether and let v be the velocity of the aether stream as it passes through the apparatus. First suppose that the arm AD of the apparatus

lies across the aether stream, and suppose that l_1 is its length in that position. To pass from A to D, the light must travel in the direction AF to allow for the drift down stream FD. Its velocity along AD is therefore

$$c\frac{AD}{AF} = c\frac{\sqrt{c^2 - v^2}}{c}$$

or $\sqrt{c^2 - v^2}$. The time t_1 for the double journey is

$$\frac{2l_1}{\sqrt{c^2 - v^2}}.$$

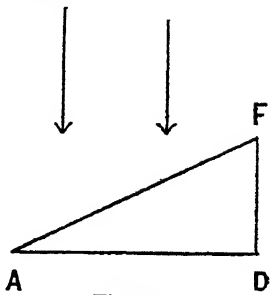


Fig. 103.

Now let AD be turned into the direction of the aether stream and have a length l_2. When travelling up-stream, the effective velocity of light will be $c - v$ and when going down-stream, $c + v$. Hence the time t_2 of the double journey is

$$\frac{l_2}{c - v} + \frac{l_2}{c + v} = \frac{2l_2 c}{c^2 - v^2}.$$

Experiment shows that the times t_1 and t_2 are equal, so

$$\frac{2l_1}{\sqrt{c^2 - v^2}} = \frac{2l_2 c}{c^2 - v^2}$$

and

$$\frac{l_2}{l_1} = \frac{c^2 - v^2}{c\sqrt{c^2 - v^2}} = \frac{\sqrt{c^2 - v^2}}{c} = \sqrt{1 - \frac{v^2}{c^2}}.$$

This quantity is less than unity; hence l_2 is less than l_1, or a body contracts in the direction of the aether stream in the ratio $\left(1 - \frac{v^2}{c^2}\right)^{\frac{1}{2}}$, where v is the relative velocity of the body and the aether, and c the constant velocity of light through the aether.

The velocity of the earth in its orbit is $\frac{1}{10000}$ the velocity of light. If, at some time of the year, this be its velocity through the aether, Michelson and Morley's apparatus would contract by one part in 200 million when turned through a right angle, and that minute change would explain their result.

There, for some years, the subject rested. Whatever the cause, every attempt to measure the velocity of light, whether with or against a supposed aether stream, led to the same result, no change in the measured velocity could be detected.

The Principle of Relativity

In 1905 an entirely new direction was given to thought on this subject by Professor A. Einstein. Einstein pointed out that the ideas of absolute space and time were figments of the imagination—metaphysical concepts, not derived directly from the observations and experiments of physics. The only space we can experience is that measured in terms of a standard unit of length, defined as the distance between two scratches on a bar, and the only time is that measured by some clock set by astronomical events. If changes such as the Fitzgerald contraction take place in our standards, they will be quite inappreciable to us who move with them and suffer corresponding changes, but they might be measurable by an observer who was moving differently. Time and space, therefore, are not absolute, but merely relative to the observer.

From this point of view the fact that the velocity of light as measured by any apparatus and in any circumstances is always the same needs no explanation. It must be accepted as the first discovered fact of the new physics. Time and space are shown by this observation to be such that light always travels relatively to us with the same measured velocity.

This measured velocity is constant, but neither space nor time nor mass measured separately show the constancy we are accustomed to expect. Michelson and Morley's apparatus, tested by our constant standard the speed of light, shows no change in linear dimensions as it rotates. But that is because we are moving with it. If we could measure accurately enough the length of a bullet as it flew past us, we should find that it appeared shorter than when at rest, and, if its speed approached that of light, it would seem very much shorter.

This experiment is not practicable; but it is easy to show, on the principle of relativity, that the mass of the bullet will

appear to an observer at rest to be increased, and increased in the same ratio as the length is shortened. This deduction may be examined experimentally. Among the marvels of modern science is the possibility of measuring the mass of projectiles which are moving past us with speeds of the same order as that of light. The β-particles, shot forth by exploding radioactive atoms, can be directed through electric and magnetic fields of force, and, as explained on page 120, their velocity and their mass can thus be determined. If the mass of a β-particle moving at moderate speeds be called unity, the following table gives in the second column the mass calculated on the principle of relativity of other β-particles, the velocity of which approaches that of light, and in the third column their mass as measured by Kauffmann experimentally.

Velocity of corpuscle in centimetres per second	Ratio of mass to the mass of a slowly moving corpuscle	
	Calculated	Observed
$2{\cdot}36 \times 10^{10}$	1·65	1·5
$2{\cdot}48 \times 10^{10}$	1·83	1·66
$2{\cdot}59 \times 10^{10}$	2·04	2·0
$2{\cdot}72 \times 10^{10}$	2·43	2·42
$2{\cdot}85 \times 10^{10}$	3·09	3·1

These β-particles are negative electrons, and, when moving, are equivalent to an electric current. Hence they create an electromagnetic field of force, which possesses both energy and inertia. The increase of mass with velocity may also be calculated on these lines with the same result. Therefore the increase of mass, like the Fitzgerald contraction, is in accordance with electromagnetic theory.

It is clear at once that these principles lead to remarkable and unexpected results. If we could travel in an air-plane (or an aether-plane) with a speed comparable with that of light, our length in the direction of motion as measured by an observer on the earth would appear to be contracted, our mass would seem greater, and our time scale slower than usual. But we ourselves should be unconscious of these changes. Our foot rule might have shrunk, but, as we should have shrunk also, we should not perceive the change. Our pound weight might have a greater mass, but so should we. Our clocks

might go slower, but the atoms of our brains would move more slowly also, and again we should not know.

But, since motion is only relative, the observer on the earth is moving relatively to us at the same rate as we are to him. Hence we should find on measuring them that his scales of length, mass and time had changed to us as ours had to him. He would seem to us to have suffered an unseemly contraction in the direction of motion, to have a mass out of proportion to his size, and to be ludicrously slow in mind and body. And all the while he is thinking the same thoughts about us. We are both unconscious of our own imperfections, but we see clearly the sad changes in the other.

It is impossible to say that either of these observers is wrong. Indeed both are right. Length, mass and time are not absolute quantities. Their true physical values are what the measurements indicate. The fact that they are not the same to everybody shows that they can only be defined relatively to one specified observer. The idea of absolute length, space, and of an absolute and even flow of time, are metaphysical concepts, which go far beyond what is indicated or justified by observation or experiment.

Nevertheless, as Bergson has pointed out, philosophically it is probable that the only time that is lived, that is, the time that measures what goes on in a system to one moving with it, or in it, is of special, indeed, of unique, importance.

But physically space and time, considered individually, are relative quantities depending on the position of the observer. It was, however, pointed out in 1908 by Minkowski that the changes in space and time compensate each other, so that a combination of the two is, even in this new world, the same for all observers. The space we are accustomed to think of has three dimensions—length, breadth and thickness, and, taught by Minkowski, we must look on time as a fourth dimension in this combination of space and time, one second corresponding to the 186,000 miles which light travels in that time. Just as the distance between two points in the continuous space of Euclidean geometry is the same however measured, so, in the new continuum of space-time, two events may be said to be separated by an "interval," involving

both space and time, which is a true absolute value whoever measures it.

In this space-time there are natural paths, like the straight path in three-dimensional space in which we think that bodies move when not acted on by a force. Since a projectile falls to the earth and the planets circle round the sun, we know that near matter these natural paths are curved, and so near matter there must be something analogous to a curvature of space-time. Perhaps, indeed, eventually we may find it possible to resolve matter into some kind of kink in space-time. However this may be, another body entering this curved region tends to move towards or round the matter in a definite path. If we prevent it from moving freely, holding it up perhaps by a chair or the surface of the ground, we exert force on it, which seems to the body to be due to its own weight.

This effect is well shown by a lift. When the lift starts upwards, it is subject to an acceleration which appears to the occupants as a temporary increase of their weight, an increase which may, indeed, be measured like ordinary weight by a spring balance. The effect of the acceleration is identical with the effect of a temporary increase in the so-called gravitational field, and it is impossible to distinguish by any experiment known to us between these two causes. This principle of equivalence, which first turned the subject towards gravitation, was set forth by Einstein in 1911, and the great mathematical difficulties were overcome during the next few years.

It then became clear that Newton's hypothesis of a gravitational attraction may be unnecessary. The movement of a body towards the earth, or round it in an orbit, may merely be the tracing of its natural path in a curved region of space-time. If we fall freely, we are conscious of no force: weight is our reaction to an upward thrust of the floor or of the platform of a weighing machine.

Calculation shows that the consequences of this theory are nearly the same as those of Newton's—quite the same to the the usual order of accuracy of observation. Yet, in one or two phenomena, it is just possible to devise a crucial experiment. Of these the most famous is the deflexion of a ray of

light by the sun, which, on Einstein's principle, is twice what the Newtonian theory would indicate. The apparent position of a star, observed near the edge of the sun during an eclipse, was found to suggest that the rays of light by which it was visible had suffered the deflexion required by the principle of relativity. To this subject we shall return in the next chapter.

Thus in two directions—in the quantum theory and in the theory of relativity—recent physics seem to be breaking away from the fundamental conceptions by which they have been guided successfully since the days of Galileo. The new thought needs new vehicles for its expression. In some ways, it is clear, the dynamics of Newton, which ushered in two glorious centuries of modern science, are proving inadequate to the tasks imposed by present knowledge. Whether after amendment they will again for a time suffice to interpret the physical universe, or whether they will pass with an honoured name into history, the discoveries of the next few years may show.

CHAPTER IX

ASTRO-PHYSICS

The History of Astronomy

THE origins of astronomy, the oldest of the physical sciences, are to be sought in the few definite results obtained by the magicians of Chaldaea and the priests of ancient Egypt.

Two thousand years before Christ astronomical results were recorded on clay tablets in Babylon. Gradually the periodicity of agricultural seasons and of astronomical events was recognized. The year was divided into 360 days, and the day subdivided into hours, minutes and seconds. The movement of the sun and the planets among the fixed stars was observed, and even eclipses of the moon successfully predicted.

Deceived, doubtless, by chance coincidences, Babylonian astrologers used this knowledge to claim the power to foretell or influence the future by observation of the stars; their services were in demand all over the known world, and Egypt, more advanced in medicine, lagged behind Babylon in astronomy.

The Chaldaeans and Egyptians pictured the world as a box, of which the earth formed the bottom and the sky the top. From it the stars hung suspended, and the sun floated round behind or above the mountains on a circumambient river.

The Greeks advanced on this crude conception. Anaximenes realized that the heavens revolved round the pole star, and saw the visible dome of the sky as half of a sphere, with the earth, a flattened cylinder, floating at the centre.

Led by the speculations of the Pythagorean philosophers, and the increasing knowledge gained by geographical discovery, Aristarchus of Samos (280–250 B.C.) put forth the view that the sun was larger than the earth, and formed the true centre of our system. But this idea was too much in advance of its time, and the authority of Aristotle helped to

postpone its acceptance till, and even after, the days of Copernicus.

Hipparchus, the inventor of trigonometry, developed the rival geocentric theory in a form preserved and amplified by Ptolemy of Alexandria (127–151 A.D.). The earth formed the centre of the world. Round it the sphere of the sky revolved, carrying the stars with it. Among the stars wandered the sun, moon and planets, moving in orbits or epicycles of their own, while these orbits themselves travelled round the earth in larger orbits or cycles. This conception of cycles and epicycles served to interpret the facts for fourteen hundred years.

For throughout the dark ages and medieval times no progress was made. First the chaos of Western Europe, and then the recovered writings of Aristotle, stopped the path. It was not until 1543 that the full work of Copernicus was printed, and a copy reached him on his death-bed.

Returning to the ideas of Aristarchus, Copernicus placed the sun at the centre of the universe, and the earth as one of his attendant planets. Beyond this system, at the confines of known space, are the fixed stars. Copernicus realized that the apparent revolution of the sky round the earth can equally well, and more simply, be represented as due to a real revolution of the earth on an axis. And on this theory, the complex arrangement of cycles and epicycles, necessary on the Ptolemaic system, was replaced by a simple movement of each planet round the sun in an orbit nearly circular.

The new ideas were verified by Galileo, who, turning his newly invented telescope to Jupiter, discovered satellites revolving round him, as Copernicus pictured the planets swinging round the sun.

Thus Galileo opened the subject of observational astronomy, as his work in dynamics opened the possibility of applying mathematically mechanical principles to its phenomena. With Galileo we enter the epoch which culminated with Newton, and held the field till the spectroscope brought new methods into use. Its characteristic triumph is the combination of mathematical astronomy based on Newtonian methods with an ever-increasing power of accurate telescopic observation.

With the application of the spectroscope and the photographic plate, we enter on the new subject of astro-physics. As Galileo and Newton taught men to apply mathematically their knowledge of terrestrial mechanics to astronomy, so Bunsen, Kirchhoff and Huggins enabled later observers to use methods and results learnt in physical and chemical laboratories to elucidate the movements and nature of the sun and stars.

Newtonian Astronomy

In the chapter on Dynamics, we saw (p. 39) that, when a body moves round a centre in a circle, we can imagine it kept in its orbit by a force F, directed towards the centre and of magnitude

$$F = \frac{mv^2}{r},$$

where m is the mass of the moving body, v its velocity and r the radius of the circular orbit. This result was proved by Huygens.

Now v is $2\pi r/T$, where T is the periodic time of one revolution. Hence

$$F = \frac{m4\pi^2 r}{T^2}.$$

It was known that, for the different planets, T^2 was proportional to r^3. Thus F must be proportional to r/r^3, that is, inversely proportional to r^2.

Led by such ideas, a young Cambridge graduate, Isaac Newton, in his rooms between the Great Gate and the Chapel in Trinity College, realized that a force directed towards the sun and inversely proportional to the square of the distance, would keep the planets in their paths if the orbits were circular, and set himself to see if it would also suffice with the actual orbits.

The facts then known were summed up in three laws, formulated by Kepler from the observations of Tycho Brahe:

(1) The planets describe ellipses about the sun in one focus;

(2) The areas swept out by the radius vector (the straight line drawn from the sun to the planet) in any orbit are proportional to the times;

(3) The sequences of the periodic times of different planets

are proportional to the cubes of their mean distances from the sun.

Newton succeeded in proving geometrically that the law of the inverse square with a force directed towards one focus would keep a body moving in an ellipse round the focus, and that no other known law would do so. He then sought for evidence that forces exist between the bodies of the solar system—a force for instance, towards the earth holding the moon in her orbit.

An apple falls from a tree. The earth pulls the apple, though we know not how. May it not also pull the moon through a longer distance with the same familiar but mysterious force of gravity? At first Newton calculated that the actual fall of the moon was too small, but six years later in 1672 an improved value for the size of the earth brought his figures into conformity, and showed that the fall of an apple and the fall of the moon were due to the same cause. Celestial phenomena were brought for the first time within range of known terrestrial principles.

The distances involved were calculated from the centres of the sun, earth and moon. Newton justified this procedure. If each particle of matter be supposed to attract every other particle with a force proportional to the product of the masses and inversely proportional to the square of the distance between them, we get for the mutual force

$$F = G\frac{m_1 m_2}{r^2},$$

an expression similar to that for electric and magnetic forces (see pages 91 and 93), except that, since no change in gravity with the medium is known, nothing analogous to the dielectric constant or the magnetic permeability appears. The quantity G, the gravitation constant, depends only on the units employed.

Taking this law of force, Newton proved mathematically that a uniform sphere of any size attracts outside bodies as though all its mass were concentrated at one point at the centre. Hence the sun and the earth, or the earth and the moon, although they are bodies of large size, attract each

other as though they were massive particles at single points.

The actual value of the minute gravitational force between two bodies has been measured in several ways. The most famous of these was used by Cavendish, and again by Vernon Boys in 1895. A light arm with two small heavy balls of gold at its ends, was hung by Boys by means of a fine quartz fibre from its middle point. Two large balls of lead are brought near the gold balls, so that their forces tend to turn the arm in the same direction. The lead balls are then swung round to the other side of the gold balls, and the difference in the two positions of equilibrium of the arm carefully measured. The couple required to turn the arm through a given angle is also determined, and thus the attractive force estimated. Boys's value for the gravitation constant is $6{\cdot}658 \times 10^{-8}$ C.G.S. units. With this value for G, the last equation gives the force in dynes between two masses measured in grams at a distance r in centimetres. From the value of gravity on the ground and the radius of the earth, its mass may be calculated as 6×10^{27} grams, with a mean density of 5·527 grams per c.c.

The complete explanation of the motions of the planets and their satellites in terms of Newton's theory is then simply a mathematical problem. By taking account of the disturbances introduced by the action of the planets on each other, nearly all the observed perturbations from their normal paths have been accounted for. One such unexplained effect suggested the presence of an undiscovered planet. Adams and Leverrier independently calculated its position, and turning his telescope to the spot indicated, Galle discovered a new planet and called it Neptune. This may be regarded as the culminating achievement of the Newtonian theory. Almost the only outstanding discrepancy was the slow shift in the line of eccentricity of the orbit of Mercury, which amounted, as measured from the earth, only to 42 seconds of arc in a century—a seemingly almost negligible error.

The Spectroscope

The origin of spectrum analysis has been described in chapter VII (page 206). Its application to examine the light from the Sun and the stars, as we saw, for the first time brought the problem of the constitution of the heavenly bodies within reach of the mind of man.

For astronomical work, the spectroscope may contain either a prism or train of prisms, or a diffraction grating (pages 197 and 207). It must be used in connexion with a telescope, and may be attached in place of the usual eyepiece, or be so placed that the light coming through the telescope can be reflected to it by a mirror. In either case, the image of the star or other object to be observed must be thrown accurately on to the slit of the spectroscope.

The spectrum so formed is usually recorded on a photographic plate, and it is convenient afterwards to throw on the same plate, on each side of the spectrum to be examined, a spectrum of some known terrestrial substance, such as iron, which gives a large number of lines. By comparison with the position of these known lines, the nature of the new spectrum and the wave lengths of its constituent light may be estimated.

Laboratory experiments with a spectroscope show four main types of spectra.

(1) Line spectra; that is, sharp bright lines on a dark ground. These are given by metals volatilized in a colourless flame, and by gases at low pressure through which an electric spark is passed. An example is given on Plate I, the frontispiece, which shows a small part of the spectrum of iron obtained from an electric arc formed between iron poles.

(2) Band spectra; broad lines or bands on a dark ground obtained from more complex bodies, or from the substances which also give line spectra, when the temperature is lower or the pressure higher. On Plate I, the lines of iron from a furnace are seen to be broader than those from the hotter electric arc.

(3) Continuous spectra; continuous bands of colour ranging from red through yellow and green to blue and violet, obtained from incandescent solids such as lime, or carbon in an electric arc.

(4) Absorption spectra; dark lines or bands on the bright ground of a continuous spectrum. These are seen most plainly in the spectrum of the Sun's light, and, in other cases, when the light from a hotter body passes through a cooler absorbing layer. As we saw on page 208, the principle of resonance shows that a vibrating system will absorb waves of the same period as it will itself emit. Hence dark lines in the Sun's spectrum, coincident with the bright lines of known elements, indicate the presence of those elements in the absorbing layer of the Sun's atmosphere. This coincidence of some few of the many lines of iron is seen on Plate I, the photograph on which covers about the one hundredth part of the whole visible spectrum.

Of these four types, the first three merge imperceptibly into each other. As the temperature is lowered or the pressure raised, sharp lines broaden out into bands, and then into a continuous spectrum.

Absorption lines, such as those in the Sun's spectrum, are really faint bright lines, which look dark by comparison with the intense light on each side of them. Their light may be seen if the more intense radiation be screened off.

Other physical influences besides temperature and pressure also have an effect. Gases when ionized (see page 123) show spectra with some lines much strengthened or "enhanced" compared with others. As ionization may be produced by high temperatures (see page 219), these enhanced lines are to be expected in the spectra of the Sun and stars, and often these lines alone are visible.

Moving electrons produce a magnetic field, and so also a magnetic field affects the radiation from electrons. Zeeman confirmed this experimentally, and showed that a strong magnetic force converted a single line into double or triple lines.

As explained on page 209, the relative approach or recession of the source of light and the observer raises or lowers the frequency of vibration, that is the colour of the light, just as the pitch of the note given by the whistle of the engine of a train is lowered as the engine passes. This Doppler effect enables us to measure the velocity of a radiating body in the line of sight.

The accuracy and delicacy of spectroscopic, and indeed of other, astronomical work has been much increased by the use of photography, especially since the application of the dry gelatine process by Sir William Huggins in 1876. In a long exposure, the effect of the light on a photographic plate is cumulative, and therefore things invisible to the eye may be made plain. The photograph gives a permanent record which may be examined and measured at leisure. It thus increases accuracy and eliminates some personal errors of observation.

We may now turn to the application of these methods and results to the chief special problems of astro-physics.

The Sun

Observations on the Sun and planets led Kepler to his descriptive laws, and enable us to construct a complete model of the solar system. The scale of the model, however, is not known until some one distance, such as that from the Sun or one of the planets to the earth, is measured in terrestrial units. From one such value, all the others may be calculated.

This fundamental measurement has been made in several ways.

(1) We have already seen on page 239 that, owing to the movement of the earth in its orbit and the finite though great velocity of light, the light from a distant star appears from the earth to be displaced from its true direction, just as a vertical shower of rain seems to come from the front of the wind-screen of a moving motor-car. Six months later, when the earth moves in the opposite direction, this displacement is reversed. These two measurements give the "aberration," and the ratio of the velocity of the earth to the velocity of light. The latter can be determined by experiments carried out on the earth, and so the velocity of the earth can be deduced from the aberration constant. Since the earth describes its orbit in a year, the size of the orbit can be calculated. Half its mean diameter gives the average distance of the Sun.

(2) When the planet Venus passes between the earth and the Sun, observation of the times of transit at two stations

gives a means of determining the distance of the Sun by trigonometry.

(3) When a planet, preferably of small size, passes near the earth, its distance may be measured directly by triangulation, observations of its position relative to a fixed star being made simultaneously at two stations on the Earth, or twice, after an interval of some hours, at the same station.

These three methods give concordant results. The value accepted at present as the most probable distance between the Sun and the Earth is 92,800,000 miles, a distance which light, moving with a velocity of 186,000 miles a second, traverses in 8·3 minutes.

The Sun's diameter can then be measured: it is about 865,000 miles. His mass may be calculated from that of the Earth, assuming Newton's inverse square law for the force between them. The mass is 332,000 times that of the Earth, and the mean density 1·4 grams per c.c. compared with 5·5 for the Earth.

The most distant known planet, Neptune, moves round the Sun in 165 years in an orbit the diameter of which is 5600 million miles. This may be considered to give the dimensions of our solar system.

As we have already said, the spectrum of the Sun shows sharp dark lines on a continuous rainbow background. The fraction shown on Plate I lies between wave lengths 5430 and 5460 tenth-metres. From the coincidence of these dark lines with the bright lines in laboratory spectra, over forty terrestrial elements have been detected in the Sun. Those which show the largest numbers of spectral lines are iron, nickel, titanium, manganese, chromium, cobalt and carbon. Others of special interest are calcium, hydrogen and helium.

Sun-spots seem first to have been observed by Galileo. Their numbers show maxima and minima, with a period of about eleven years. When there are most spots, the Sun's radiation is most intense, and magnetic disturbances are most violent on the earth.

By watching sun-spots it may be seen that the equatorial regions of the Sun revolve once in about 25 days, the speed of rotation being greater there than it is nearer the poles.

This is confirmed by turning a spectroscope first to one edge of the Sun's disc and then to the opposite edge. Some of the spectral lines are displaced, in accordance with the Doppler effect. The east limb of the Sun is approaching and the west limb is receding. Other lines show no displacement. These lines, therefore, cannot be due to the Sun, but are caused by absorption in the Earth's atmosphere.

The spectrum of the light from a sun-spot is not identical with that from the rest of the Sun. The lines are broader, as seen in one of the photographs on Plate I, and some of them indicate the presence of compounds such as titanium oxide, thus suggesting a lower temperature.

By cutting off the light of the surrounding spectrum, the Sun may be photographed by the light of one spectral line if it be intense enough. Lines due to calcium and hydrogen have been used, and the photographs obtained show glowing clouds of these elements covering the surface of the Sun.

Measurements of the heat absorbed by a calorimeter exposed to the Sun's rays at high altitudes, enable us to determine the rate of heat emission, a correction being made for the heat stopped by the earth's atmosphere. It seems that each square centimetre of the Sun's surface emits 89,000 gram-Centigrade thermal units per minute. This corresponds to about nine horse power.

Both thermodynamic theory and experiment show that the total radiation from a body which emits a continuous spectrum is proportional to the fourth power of the absolute temperature. Hence from the Sun's radiation his temperature may be calculated. The figures given indicate that the radiating layer is between 5000° and 6000° C. Inside the Sun, the temperatures must of course be very much higher.

Outside the radiating layer lies the absorbing layer, still luminous, but so much less so that it absorbs more light than it emits, and thus gives rise to the characteristic solar spectrum with its dark absorption lines.

Beyond that is another cooler layer, the chromosphere, seen just before an eclipse becomes total, and occasionally tossed up into gigantic prominences. Its spectrum may be seen momentarily at eclipses, or may be photographed at

any time by aiming the spectroscope tangentially at the Sun's extreme edge. It consists of bright lines, some of which, those of helium, for example, are not found in the Fraunhofer spectrum. When an eclipse is total, and then only, another phenomenon, the Corona, may be seen. It consists of streamers of faint light reaching to a great distance from the Sun.

The source of the Sun's enormous output of energy is a difficult problem. As the matter which forms the Sun fell together, it would convert its potential energy into heat, and shrinkage may still be going on. If radioactive bodies are present, as the existence of helium suggests, they may supply some of the energy. But it is probable that both these effects are insufficient either alone or together, and some other cause must be at work. We should remember that we have no experience on earth of such tremendous temperatures as those in the Sun. Changes quite unknown to us may be going on. The energy may possibly be produced by atomic disintegration, or by some even more unfamiliar cause, such as a direct conversion of colliding electrons into radiation.

The Stars

Beyond the outer confines of the solar system lies a great gulf of space. By careful observation, the nearer stars may be seen to move relatively to those more distant, as the Earth passes in six months from one side of its orbit to the other. Another six months brings them back again, save for any small shift due to the real movements of the stars themselves. Corrected for this latter change and for the aberration of light, a star's six-months' parallax gives its distance by triangulation, since we know the diameter of the earth's orbit.

The nearest star, a faint speck called Proxima Centauri, is thus found to be 24 million million ($2{\cdot}4 \times 10^{12}$) miles away from us—a distance, four thousand times the diameter of Neptune's orbit, traversed by light in 4·1 years. Sirius, the bright dog-star, is 5×10^{12} miles or 8·6 light-years away.

On a clear night, the eye may see a few thousand stars. Others become visible as we use telescopes more and more powerful, but the number revealed does not increase in proportion to the power of the instrument, and thus it may be

concluded that the total number is not infinite. The large 100-inch reflector of the Mount Wilson Observatory in America shows a number estimated at 100 million, and it would seem that in the stellar universe the total is some such number as 1500 million.

The brightness of the stars varies greatly, and this gave the earliest method of classification. Hipparchus called the twenty brightest, stars of the first magnitude, and those he could just see he grouped in the sixth magnitude. The relative brightness of first and sixth magnitude stars is about 100 to 1, and, if we take these exact figures and arrange magnitudes so that the brightnesses are in geometrical progression, we get $100^{\frac{1}{5}}$ or 2·512 as the ratio of the brightness of stars of one magnitude to that of stars of the next lower magnitude. This scale is continued to stars only visible in telescopes. Thus the brightness of stars of the sixteenth magnitude is $1/(2{\cdot}512)^{15}$ or about the one millionth part of the brightness of stars of the first magnitude.

This system depends, of course, on the apparent brightness of the stars as seen from the Earth. For those stars whose distance is known, we can calculate the apparent magnitude which the star would have if moved to a standard distance, and call it the absolute magnitude.

When classed according to their absolute magnitudes, we find stars of all values, but more in the higher and lower than in the intervening magnitudes. These are called "Giant" and "Dwarf" stars respectively. They will be dealt with more fully below.

Stars of the same spectral type show a regular relation between absolute magnitude and the relative intensity of certain spectral lines. Hence a careful examination of these critical lines gives a value for the absolute magnitude of a star. Then, from its apparent magnitude, its distance may be estimated, even when it is too great to be measured by parallax. This gives one of several indirect methods of estimating stellar distances.

Double and Multiple Stars

Many stars which look single to the naked eye are seen to be double through a telescope. The individual stars in some of these

pairs may be far from each other, and only seem near because they are almost in the same line of sight. But the number of double stars is much too great for such chance conjunctions to explain them all. In most cases, there must be a physical connexion.

The law of gravitation indicates that each component of a binary system must describe an elliptical orbit about the common centre of gravity as one focus. And, for one of those few double stars in which the orbits and distances from us have been observed, it is possible to calculate the total mass of the pair. These masses are found to range from about half to about three times the mass of the Sun, in agreement with other evidence which shows that the difference in mass between different stars is not very great.

Some double stars are too near each other to be separated by a telescope, but can be resolved spectroscopically. If their orbits are seen edge-on, and the line joining the two stars is perpendicular to the line of sight, one star will be approaching and one receding from us. Hence, by Doppler's principle, the lines in one spectrum will be shifted towards the blue and those in the other towards the red, and, in the actual spectrum of the double star, the lines will be doubled. When the stars are one behind the other, they will be moving across the line of sight and no spectral doubling will appear. By observing these changes in the spectra, the period of revolution and the velocities may be estimated, and the ratio of the masses calculated. Hence, if visual as well as spectroscopic measurements are possible, the individual masses can be determined.

The light of many stars varies in strength from time to time. When it varies irregularly, the changes may possibly be due to recurrent outrushes of incandescent gas, but very often the period of change is quite regular, and the cause of variation may be referred to the eclipse of the bright star by an invisible companion, which cuts off some or all of the light. Sometimes this interpretation can be confirmed spectroscopically, the spectral lines being displaced periodically when the bright star is approaching or receding radially from the Earth. From a curve of variation of light with time, combined with spectrum measurements, a very complete description of

the system can often be obtained, as, for instance, with the stars named Algol and β Lyrae.

Other regular variables, such as δ Cephei, cannot easily be explained by eclipses. When of short period, these "Cepheid" stars show a definite relation between the period of variation and the luminosity or absolute magnitude. This is so regular that measurement of the period of other similar stars at unknown distances has been used as a means of estimating the absolute magnitude. An observation of the apparent magnitude then gives the distance—a second method applicable to stars too far away to show any parallax.

The number of double stars is immense, and still more complex systems, multiple stars, can be recognized and examined by the same methods. For example, the well-known "Pole Star" has been found spectroscopically to be a triple star with periods of four days and twelve years, revolving round a fourth visible star in some such time as twenty thousand years.

The Spectra of Stars

The most recent system of classification of stars depends on their spectra. It was begun by Father Secchi, and has been much improved and extended at Harvard Observatory.

Even the visual colours of stars differ. Since photography is more sensitive to the violet end of the spectrum, photographic magnitudes are not the same as those estimated by eye, and the difference between them gives a scale of colour.

These differences are further exemplified in the spectra. A series of spectral types can be found, passing into each other by insensible gradations, but showing definite characteristics which have been distinguished at Harvard by the letters O, B, A, F, G, K, M, N, R, a list in which the bluer stars come first.

Type O spectra show a faint continuous background on which appear bright lines. In some, the hydrogen and helium lines are prominent.

Type B spectra give dark lines with helium very prominent. Type A show hydrogen, and also calcium and other metals, which increase in importance in type F.

Type G includes our Sun. The spectra show dark lines on a bright ground. The colour is yellow. In type K, bands due to hydrocarbons appear for the first time.

Type M stars give broad absorption bands, especially those of titanium oxide. Type N spectra show broad absorption lines due to carbon monoxide and cyanogen. These stars are red in colour. Type R stars also show the absorption bands found in N, though the colour is not so red.

These observations on spectra enable us to estimate the effective temperatures of different types of stars.

If a black body, which may be regarded as a perfect radiator, is gradually heated, the character as well as the intensity of the radiation changes, as we have already seen on page 211. For each temperature there is a characteristic curve between radiant energy and wave length, showing a maximum at some particular wave length. The position of this maximum shifts towards the blue end of the spectrum as the temperature rises and thus indicates the temperature. The distribution of energy has been investigated in several ways, both photographically and by studying variations in the character of the radiation. Direct relations between temperature and the appearance of certain spectral lines have also been used to estimate the stellar temperatures.

Stars of type R are the coolest which are visible. They have effective temperatures of about 2300° C. From R we pass upwards through G, about 5000°, till we reach some 23,000° C. in the hottest stars of type O.

These temperatures are of course those of the radiating layer; the inside parts of a star must be much hotter, with temperatures rising, perhaps, to a million degrees.

When considering the absolute magnitudes of stars, we saw that, while intermediate stars were not unknown, they mostly fell into two groups, "giants" and "dwarfs," one with much greater luminosity than the other. Now it is remarkable that this division into groups is only clear in the cooler stars of types K onwards, with temperatures not above about 4000° C. In the hotter stars, the division is less marked, and in type B it has quite disappeared; these stars are all "giants," with luminosities from 40 to 1600 times that of our Sun.

These facts point to a definite conclusion, namely, that stars go through a course of evolution roughly identical. They begin comparatively cool; gradually they rise in temperature; they attain a maximum depending on the size of the star; and then pass down the same scale as they cool again.

While ascending in the scale, a star emits a very large amount of light and is of high absolute magnitude. This means that it must be of enormous size. It is classed as a "giant star." As it cools, its atmosphere passes through the same range of temperature in the reverse order, and so the star goes through the same spectral types in its descent as in its rise, though certain differences in detail are observed. But now its absolute magnitude, that is, its luminosity, is much less, which, since the temperature is the same as in its rise, shows that it is much smaller in size. The star has become a "dwarf."

This process of stellar evolution, traced by Russell, is in accordance with the dynamics of a mass of gravitating gas, as worked out by Lane and Ritter. If the mass be large enough, gravity will cause it to contract. It will give out heat and grow hotter. But, as it shrinks, the rate of contraction must decrease. At a certain critical density, the heat developed by this colossal mass of glowing vapour becomes less than that radiated, and the mass begins to cool. As we saw when considering the Sun, this process may not explain all the heat evolved, but it is probable that other sources of energy, such as atomic disintegration, may depend on temperature and go through a similar history.

We seem to have traced the normal course of stellar evolution. Is there any direct evidence for this theory in the depths of space?

Nebulae

Even in the largest telescope a star has no visible dimensions —the nearest is too far away. But areas of luminosity, called nebulae, have long been known. One, the great Nebula in the constellation Andromeda, being visible to the naked eye, was observed before the invention of the telescope, and another situated in Orion was discovered by Huygens in 1656.

There are three chief classes of nebulae:

(1) irregularly shaped nebulae such as that in Orion;

(2) planetary nebulae, smaller bodies of regular form;

(3) spiral nebulae, like the "Whirlpool" shown on Plate I.

Irregular nebulae. Not only are there bright bodies of irregular form, like the nebula in Orion, but also large areas in the sky of faint nebulosity, passing into dark nebulae which blot out the more distant stars in certain regions of star-spangled space.

The characteristic spectra of the bright, irregular nebulae, consist of bright lines on a faint continuous background. The lines of hydrogen and helium are prominent, together with other unknown lines, referred to a hypothetical element called "nebulium."

The Orion nebula as a whole is not moving fast either towards or away from us, but determinations of radial velocity show great local movement, and suggest also a general rotatory motion.

Planetary nebulae. This name is given to small nebulae of regular outline, which in a telescope look of sensible size, and resemble in appearance the outer planets. When their shape is elliptical, as in the smaller nebula shown on Plate I, they seem to rotate about the shorter axis.

The spectra are of the same general type as those of stars which are classed in the Harvard group O and show bright bands. Indeed there seems no sharp division to be drawn between stars of type O and planetary nebulae.

Spiral nebulae. The greatest number of nebulae are of spiral form, such as the larger on Plate I. Their spectra are continuous, with superposed absorption lines—spectra which resemble those of stars of classes F to K, including our Sun. The lines are in general displaced in the direction which seems to indicate that the nebulae are moving away from us with high velocities.

They also show evidence of rapid rotation. Nebulae which we see edge-on, can be examined spectroscopically, and some of those which are at right angles to us, when photographed year after year, also show measurable rotation, indicating a rate of one revolution in some hundred thousand years. This

looks slow, but linear velocities up to a thousand kilometres a second have been observed, so that the long period of rotation is due to colossal size rather than to slow movement.

If we assume that different nebulae rotate with velocities of the same order, a comparison of the radial speed of edge-on nebulae measured spectroscopically with the annual angular rotation of nebulae with planes across the line of sight gives an estimate of distance. From this and other evidence, values have been obtained ranging from tens to hundreds of thousands of light-years. The spiral nebulae are very far away.

The nebular theory of evolution was first suggested by Laplace at the end of the eighteenth century, in an attempt to explain the origin of the solar system. Laplace began with the idea of a gaseous nebula, which filled the space comprised in Neptune's orbit, and possessed a motion of rotation. Under its own gravitation, the nebula contracted, and therefore, since the angular momentum was constant, it moved with increasing velocity. At various stages, it left behind as it shrank rings of matter, which condensed into the planets and their satellites, revolving round the central mass which formed the Sun.

Now there are several difficulties in this interpretation. For instance, mathematical investigation shows that a mass of gas of the dimensions and density required would not contract. Its gravitation is not powerful enough to overcome this diffusive effect of gaseous pressure, and the nebula would be scattered through space.

But in the spiral nebulae we have bodies a million times larger than that imagined by Laplace, and on this scale the whole course of change is different. Gravitation is now more effective than both gas and radiation pressure, and the nebula, instead of dissolving, contracts and spins faster as Laplace supposed. The explanation fails for the comparatively small solar system, but may succeed for a gigantic stellar galaxy.

Jeans has proved mathematically that a mass of gravitating gas, set in rotation perhaps by the tidal action of other masses, will gradually assume the form of a double convex lens, such

as is seen in the smaller nebula of Plate I. As it spins faster, the edge must become unstable, and break up into arms like those of the spiral nebulae. Local condensations will occur in the arms, each of the appropriate size to form a star within the somewhat narrow limits of size we find stars to possess. In spiral nebulae, then, we see stellar systems in the making, it may be within our visible system of stars, it may be beyond, where other systems may be forming.

Will one tiny globule on the arm of a spiral nebula form a solar system such as ours? Mathematics indicate that it is not probable. If the rotation of the globule be fast enough to cause disruption, it seems that a double star with two partners waltzing round each other should result. Thus the multitude of double stars probably show the alternative normal course of stellar life to that which leaves a solitary single star.

Jeans has offered a speculative theory to account for our own system. If, at an early period, two stars came near each other, tidal waves would appear. If they approached within a certain critical distance, such a wave would shoot out a long arm of matter, which might break up into bodies of appropriate size and character to form the Earth and other planets. But this would rarely happen, and our system may be unique or nearly so.

The Stellar Universe

We have now described the chief methods and some of the detailed results of astro-physical research, and may collect the information to obtain some idea of the general structure of the stellar Universe.

Stars are most numerous in a band of varying width, called the Galaxy or Milky Way, which stretches round the heavens nearly in a great circle. In places, the numbers are so great that they appear as star-clouds, only to be resolved into individual stars by good telescopes, while interspersed are irregular nebulae which cannot be resolved.

The great plane, which cuts the Milky Way as near as may be in the middle of the band of stars, is called the galactic plane. It may be looked on as a plane of symmetry

in the stellar system. Towards it the stars seem to crowd, especially the hotter types B and A, and the fainter stars, that is, those which, on the average, are farther away.

This indicates that the stellar Universe is flattened in the galactic plane. It seems to form a vast lens-shaped system. We are within it, but not at the centre. We see more stars in the Milky Way, chiefly because, when looking at it, we are looking towards the edge of the lens, where the depth of star-strewed space is much greater than elsewhere.

Besides star-clouds and irregular nebulae, about a hundred globular clusters of stars are also known. They are most numerous just outside the central zone of the Milky Way. They contain Cepheid variables, and, from the period of these and by other indirect methods, their distances have been estimated to range from 20,000 to 200,000 light-years.

Thus it appears that our stellar system has a longest diameter which it would take light at least 300,000 years to traverse. Our Sun lies about 60,000 light-years from the centre of the whole system, somewhat to the north of the central plane. Observation over many years of the apparent movements of the stars shows that the Sun is moving towards the constellation called Hercules with a speed of about 13 miles a second, and that, taking this drift as a line of reference, there are two main directions in which the stars stream through space.

Unlike other heavenly bodies, spiral nebulae are never found in or near the Milky Way. While some appear to be only about 25,000 light-years away, the distances of others are immense, estimates of 500,000 light-years having been made. It is possible that, while some are forming new star-clusters within our stellar system, others, beyond its confines, are giving birth to new systems. We must suspend our judgment pending further evidence.

The new outlook on nature given by the theory of Relativity, as described in the last chapter, must, as it unfolds, affect profoundly our views about the physical Universe.

It is true that, to a high order of approximation, Newton's theory of gravitational attraction is still in accordance with

facts, and can be used to deduce the phenomena of planetary and stellar movement and evolution. It is only with difficulty that crucial experiments have been devised to decide between Newton and Einstein. One such has been already referred to on page 248. On the principle of relativity, light is deflected from its path by a gravitational field to twice the extent to be expected on the older theory. The only way in which these minute deflexions can be observed is to photograph during an eclipse of the Sun the image of a star which appears just outside the Sun's disc. This was done by Eddington at Principe in the Gulf of Guinea and by Crommelin in Brazil during the eclipse of 1919. Compared with stars further away from the Sun it was found that the image of the near star was displaced, and displaced to the amount required by Einstein.

Secondly, the discrepancy of 42 seconds of arc per century in the orbit of Mercury, left over by the Newtonian theory, was at once explained by Einstein, who calculated a change of 43 seconds of arc.

Thirdly, on the principle of relativity, an atom should vibrate more slowly in a gravitational field. Hence, on the average, the lines in the spectrum of the Sun, where gravity is more intense, should be displaced towards the red compared with the lines in corresponding terrestrial spectra. The shift to be expected is barely perceptible, but the balance of experimental evidence is now in its favour.

Thus it seems that, as an exact account, Newton's theory must give place to Einstein's. But the replacement of the idea of attractive force as an ultimate explanation by the theory of natural tracks, which appear to us curved in gravitational fields, not only leads to slightly different results in the exact experiments just described, but must change completely our ideas about the confines of the Universe.

With Euclidean space and Newtonian time, we naturally thought of the Universe as infinite. Space stretched indefinitely beyond the farthest stars, and time, both before and after us, flowed on uniform and eternal.

But, if our new space-time continuum be curved, we enter on another range of thought. Time may still run from ever-

lasting to everlasting in a never ending series of moments, but the curvature of space means a Universe finite in its space dimensions. If we travel long enough with a ray of light we shall meet a limit, or perhaps return to our starting point.

Beyond that is speculation. Einstein's three-dimensional space is curved in a way which, in two dimensions, we should call cylindrical, and thus, by travelling outwards for ever, we should return to the point from which we set forth. Time is like the axis of the cylinder. De Sitter imagines a spherical space-time. If we travel outwards, and trace ever wider spheres, we come to one of maximum size. Here time observed from earth seems to stand still. As Eddington says: "Like the Mad Hatter's tea party, it is always six o'clock and nothing whatever can happen however long we wait." But, if we could get to this sphere we should find time really running on, though in another direction—whatever that may mean.

One slight indication of such a slowing down of time has been pointed out by Eddington. Some of the spiral nebulae are the most remote bodies known to us. Their spectral lines are displaced compared with the same lines in terrestrial spectra, and, in a large preponderance of cases at least, the displacements are towards the red. This is usually interpreted as due to very great velocities of recession—greater than any others observed in celestial bodies. But it is just possible that we are here observing the slowing down of atomic vibration as seen from the Earth, a change in the rate of Nature's clock, a variation in the scale of time. Once more we must suspend our judgment, and wait for further light.

We have finished our survey of the present state of knowledge of the physical Universe. We may well be amazed both at the extent of human achievement and at the veil of ever-growing mystery beyond which lies the meaning of it all.

INDEX

α rays, 227, 228, 231, 234
Aberration of light, 239, 256
Absolute scale of temperature, 57–59, 74–76
Acceleration, 4, 25, 26, 30, 31
Accessory food factors, 178
Accumulator, 99
Acetaldehyde, 163
Acetic acid, 163
Acetylene, 160, 165
Acid forming elements, 15–18
Acidity, 143
Acids, acetic, 163; aromatic, 166; benzoic, 166; butyric, 164; formic, 163; lactic, 158, 159; organic, 163; racemic, 157; salicylic, 166; strength of, 143; tartaric, 157
Actinium, 226
Adams, J. C., Neptune, 253
Aether, the luminiferous, 238
Air, conductivity of, 119; resistance, 2, 31; thermometer, 53–57
Alcohol, ethyl, 155, 161; melissyl, 162; methyl, 161
Aldehydes, 162
Algol, 262
Alloys, 131–133
Alternating currents, 110, 112
Aluminium, 14
Amber, 90
Amino-acids, 176, 179, 180
Ammeters, 105
Ampère, A. M., combination of gases, 11
Ampere, the, unit of current, 100
Amplitude of vibration, 184
Anaximenes, astronomy, 249
Andrews, T., liquefaction, 82–85
Anthracene, 165
Archimedes, floating bodies, 2, 42
Argon, 73
Aristarchus, astronomy, 249
Aristotle, astronomy, 249; falling bodies, 1, 6
Aromatic acids, 166
Aston, F. W., isotopes, 20, 122, 231; positive rays, 121
Astronomy, history of, 249
Astro-physics, 249
Asymmetric carbon atom, 158, 170
Atmosphere, pressure of, 46
Atom, the individual, 231
Atomic heat, 67, 74; numbers, 21, 22, 225, 230, 233; radiation, 234; structure, 232, 237, 238; theory of electricity, 118; weights, 11, 12, 20, 21, 22, 45, 230
Atoms and molecules, 10, 45, 72, 154 *et seq.*, 231, 237
Aurora borealis, 123
Available energy, 77, 124
Avogadro, combination of gases, 11, 12; hypothesis, 11, 12, 73

β rays, 227, 245
Babylon, 249
Barometer, 46
Becquerel, H., radioactivity, 226
Bell, Graham, telephone, 108
Benzene, 164, 179
Benzoic acid, 166
Bergson, H., time, 246
Biochemistry, 151, 167
Blood, neutrality of, 174
Bohr, N., atomic structure, 237, 238
Boiling point, 78, 79
Boyle's law, 47, 72, 73, 86
Boys, V., gravitation constant, 253
Bradley, J., aberration of light, 239
Bragg, Sir William, X-rays, 224
Bragg, W. L., X-rays, 224
Breakers, 186
Buffer substances, 174
Bunsen, R. W., astro-physics, 251; spectra, 208
Burning, 6, 7, 152
Butane, 155, 160

Butylene, 159, 160
Butyric acid, 164

Cadmium, crystals of, 132
Calcium, 14
Caloric fluid, 59
Calorimeter, 60, 61, 65, 67, 68
Calorimetry, 60
Canal waves, 184
Cane sugar, 169
Capacity, electric, 215
Capillarity, 50
Carbides, 159, 160
Carbohydrates, 167, 169
Carbolic acid, 165
Carbon and its compounds, 19, 20, 151; dioxide, 19, 179; liquefaction of, 82–85; monoxide, 19, 150
Carbon atom, asymmetric, 158
Carnot, S., heat engines, 74, 81
Casein, 176
Catalysis, 149, 167
Cathode rays, 119–121, 223
Cattle cake, 173
Cavendish, H., electric forces, 91; gravitational constant, 253; hydrogen, 9; pneumatic trough, 3; water, 9
Cell, accumulator, 99; voltaic, 98
Cellulose, 169, 171
Celsius, A., centigrade scale, 53
Centigrade scale, 53
Centrifugal or centripetal force, 98
Cepheid stars, 262, 268
Chaldea, 249
Charcoal, 19
Chemical action, 6, 146; equilibrium, 124, 143
Chemical energy, electric equivalent of, 101
Chemical equivalent weights, 11, 117
Chemical potential, 126
Chloral hydrate, 162
Chlorides, 17
Chlorine, 16, 17, 45; density of, 45; liquefaction of, 17; manufacture of, 147
Chlorobenzene, 165
Chloroform, 154
Chlorophyll, 179
Circular motion, 37, 38
Clerk Maxwell, daemon, 77; dielectric medium, 122; electromagnetic waves, 115, 213
Cloud or mist, 52
Cloud tracks, 232
Coagulation of colloids, 134
Coal, distillation of, 160, 164, 166, 180
Colloids, 49, 134
Combination, 6, 7; apparatus, 152
Condenser, electric, 113
Conduction of electricity through gases, 118; liquids, 116; solids, 123
Conservation, of energy, 36; of matter, 9, 65, 245
Constitutional formulae, 154
Copernicus, N., astronomy, 250
Copper, 13–15, 131, 132
Corpuscle, electric, 121, 123
Correlation of forces, 63
Coulomb, C. A., electric forces, 90
Couples, 36
Critical point, 84
Crommelin, A. C., deflexion of light, 269
Crookes, Sir William, radium, 231
Cryohydrates, 130
Crystal receiver, 219
Crystalloids, 134
Crystals, 12; of metals, 2; structure of, 224
Curie, M. et Mme, radium, 226, 229

Dalton, J., atomic theory, 10–13
Daniell's cell, 99
Davy, Sir Humphry, potassium and sodium, discovery of, 21, 99; heat of friction, 63
Decane, 159
Declination, magnetic, 59
Deflexion, magnetic and electric, 120
Degrees of freedom, 125, 127
Democritus, atomic theory, 10
Density, of gases, 43; of vapours 44
Density and specific gravity, 42
De Sitter, space-time, 270
Dewar, Sir James, liquefaction of gases, 86

Dextrose, 169
Diamond, 19
Diastose, 171
Dielectric constant, 91, 92, 115, 213
Difference of potential or electromotive force, 101
Diffraction grating, 197
Diffusion of gases, 72
Dimensional equations, 29, 52, 185
Dimolecular reaction, 148
Dispersion, 206
Dissipation of energy, 76
Dissociation of hydrogen selenide, 144; of iodine vapour, 146; of nitrogen peroxide, 147; of water vapour, 145
Dissociation theory, 141
Doppler effect, 209, 257, 258, 261
Double stars, 260; spectroscopic, 261
Dufay, electricity, 90
Dwarf stars, 260, 263
Dyes, 166
Dynamical equivalent of heat, 63, 64, 105
Dynamics and the physical properties of matter, 23
Dynamo, 109–111
Dyne, the, unit of force, 29

Earth, coil, 109; quadrant, 23
Earth, the, density of, 253; magnetism of, 95, 97, 98, 100; mass of, 253
Eddington, A. S., deflexion of light, 269; space-time, 270
Egypt, 249
Einstein, A., gravitation, 6, 247; quantum theory, 236; relativity, 243, 269, 270
Elasticity, 41; of a gas, 47; of shape, 2, 41; of volume, 2, 41
Electrical thermometer, 57–59, 104
Electric capacity, 92, 215; charges, 91, 92; condenser, 113; current, 98; deflexion, 120; fish, 239; force, 90, 91; induction, 91; inertia, 233; lines of force, 216; machine, 91; oscillations, 214, 219, 220; potential, 92; power, 111; resistance, 102; supply, 110; telegraph, 101; units, 114; ratio of units, 115
Electricity and heat, 104
Electricity and magnetism, 90
Electricity, atomic nature of, 118
Electrochemical action, 99; equivalent, 117, 121
Electrolysis, 116, 117, 122
Electrolytes, conductivity of, 118, 146; ionization of, 141–143; osmotic pressure of, 140
Electrolytic polarization, 98
Electromagnetic induction, 107, 108; machinery, 109; waves, 115, 213, 214, 216, 238
Electromotive force, 101; origin of, in cell, 101
Electro-plating, 99
Electronic theory, 122
Electrons, 121, 123, 232, 238
Electroscope, 91, 92, 101, 118
Electrostatics, 90
Elements, 6; table of, 22; undiscovered, 21, 225, 227
Elizabeth, Queen, 90
Emission theory of light, 198
Empirical formulae, 153
Energy, 29, 30, 34; available, 77, 124; conservation of, 36, 62–65; dissipation of, 76; path of, 123
Engine, heat, 74–76; internal combustion, 76; steam, 76
Enzymes, 167, 168
Equations, gaseous, 86–89
Equilibrium, 30; chemical, 143, 171; of phases, 61, 62, 80
Erg, unit of work, 29
Esters, or ethereal salts, 164
Ethane, 155, 160
Ethylene, 10, 11, 160
Euclid, space, 25, 269
Eutectic alloys, 132
Evaporation, 51, 81, 125
External work of gases, 69–71, 85

Fahrenheit scale, 53, 54
Falling bodies, 1, 30, 31
Faraday, Michael, dielectrics, 91; dielectric medium, 122; electrolysis, 116, 118; electromagnetic induction, 107, 108; induction

of currents, 214; ions, 117; quantity of electricity, 91
Fatigue, 172
Fats, 167, 178
Fermentation by yeast, 171
FitzGerald, G. F., contraction, 242, 245
Fizeau, velocity of light, 240
Fluids, 2, 45, 51
Flux, magnetic, 108, 109
Focal length, 201, 205
Force, 28, 29; idea of, 4
Formaldehyde, 162, 170, 179
Formalin, 163
Formic acid, 163
Formulae, constitutional, 154, 160; empirical, 153; molecular, 154; ring, 165
Foucault, L., spectra, 208; velocity of light, 240
Frankland, Sir E., helium, 208
Franklin, Benjamin, 90
Fraunhofer, J. von, spectral lines, 207, 255, 257, 259
Free expansion, 69
Freedom, degrees of, 125, 127
Freezing point, 129; of solutions, 138, 139; of water, 53, 127, 129
Frequency of waves, 184
Fresnel, A. J., interference, 196
Fructose, 169
Fruit sugar, 169

γ rays, 227
Galactose, 169
Galaxy, the, 267
Galileo, satellites of Jupiter, 250; sun spots, 257
Galle, Neptune, 253
Galvani, electric currents, 98
Galvanometer, 99–101
Gas, constant, 72; equations, 86–89; laws for solutions, 136
Gases, 3; combination of, 11, 12; conductivity of, 118; density of, 43; diffusion of, 72; external work of, 69; free expansion of, 69; internal work of, 69; kinetic theory of, 71–74; and liquids, 82; liquefaction of, 85, 86; monatomic, 73; ratio of, 68, 69; specific heats of, 65–71
Gay Lussac, combination of gases, 11; free expansion, 69
Giant stars, 260, 263
Gibbs, Willard, one-component systems, 127
Gilbert, W., electricity, 90
Gilbert, Sir John, feeding experiments, 174
Glucose, 169
Glycerine or glycerol, 162, 164
Glycine, 176–178
Glycol, 162
Glycogen, 170, 172
Gold, 13, 15; mines, 15
Goldstein, positive rays, 121
Graham's law, 72
Gram-molecule, 45, 67, 87
Grape sugar, 169
Grating, diffraction, 197
Gravitation, 5, 27, 247, 251, 268–270; constant, 252, 253
Gray, Stephen, electric conductors, 90
Gun cotton, 172
Guthrie, cryohydrates, 130

Haber, nitrogen, 18, 180
Haemoglobin, 180
Half-period elements, 194
Harmonic, motion, 38; oscillations, 37, 214
Harvard Observatory, 262
Heat, 53; and electricity, 104; and energy, 53; engines and the absolute scale of temperature, 74–76, 81; latent, 61, 62; quantity of, 59; unit of, 59, 60
Helium, 208, 228, 233
Hertz, H., electromagnetic waves, 214
Hipparchus, astronomy, 250
Hopkins, F. G., diet, 178
Horse-power, 29
Huggins, Sir William, astrophysics, 251; photography, 256
Huygens, C., centripetal force, 251; nebula in Orion, 264; refraction of waves, 201
Hydrocarbons, saturated, 159; unsaturated, 160
Hydrochloric acid, 16

Hydrogen, 11, 12, 45; atom, 233, 238; chloride, 11, 12, 16, 45; electrode, 143; ion concentration, 143, 150, 174; molecular weight of, 45; nascent, 150; nucleus, 233, 234; and oxygen, 145; porous plug effect, 70; preparation of, 8; and selenium, 144; thermometer, 54, 57
Hydrolysis of fats, 164, 168
Hydrostatics, 42
Hydroxyl, 150, 156; groups, 162

Ice, melting point of, 129; water and steam, 3, 127
Iceland spar, 212
Images, optical, 199, 204, 205, 206
Impact, 28
Index of refraction, 203
Indicators, action of, 19
Induction coil, 112
Inertia, moment of, 36; and mass, 3, 9
Interference, of light, 195; of sounds, 192, 194
Internal work of gases, 69–71, 85
Interval in space-time, 246
Invertase, 169
Iodine vapour, dissociation of, 146
Ionization, 117, 123, 141, 174; of electrolytes, 141–143, 175
Ions, electrolytic, 117; recombination of, 149; velocity of, 118
Iron, 14, 15; crystals of, 132
Isobutane, 155
Isomers, 155; physical, 156, 158
Isotopes, 20, 122, 225, 231

Janssen, P. C. C., helium, 208
Jeans, J. H., nebular theory, 266; origin of solar system, 267
Joly, M., steam calorimeter, 61, 65, 67, 68
Joule and Thomson, porous plug experiment, 70
Joule, J. P., dynamical equivalent of heat, 63, 64; electric heat effect, 104; free expansion, 69; kinetic theory, 13, 71–74

Kauffmann, β rays, 245
Kelvin and Joule, porous plug experiment, 70
Kelvin, Lord (W. Thomson), absolute scale of temperature, 75
Kepler, J., planetary laws, 251
Kinetic, energy, 35; theory, 13, 59, 71–74
Kirchhoff, G. R., spectra, 208; astro-physics, 251
Kjeldahl, estimation of nitrogen, 153

Laborde, radium, 229
Lactic acid, 158, 159
Lactose, 169
Laevulose, 169
Lane, H., stellar evolution, 264
Laplace, Marquis de, nebular theory, 266; velocity of sound, 188
Larmor, Sir Joseph, electrons, 232; Fitzgerald's contraction, 242
Latent heat, 61, 62; equation, 81, 128, 138
Laue, X-rays, 224
Lavoisier, A. L., carbon compounds, 151; conservation of matter, 26; hydrogen, 9; oxygen, 7, 8
Lawes, Sir John, feeding experiments, 174
Lead cell or accumulator, 99
Lead from radium, 122, 231
Le Bel, rotatory power, 157
Lenses, 204
Leucine, 177
Leucippus, atomic theory, 10
Lever, the, 34
Leverrier, Neptune, 253
"Life," 181
Light, aberration of, 239, 256; composition of, 206; deflexion of, by gravity, 248, 269; emission theory, 198; interference of, 195; polarization of, 211; reflexion of, 198; refraction of, 201; velocity of, 115, 116, 213, 239–244; wave theory of, 195, 197, 211
Lines of electric force, 94, 216
Linnaeus, C., centigrade scale, 53

Lipase, 168
Liquefaction of gases, 85, 86
Liquid and gaseous states, relations between, 82
Liquids, 3, 42, 45, 49, 51, 82
Litre, 24
Lockyer, Sir Norman, helium, 208
Lode stone, the, 93
Lodge, Sir Oliver, velocity of light, 242
Lorentz, electrons, 232; Fitzgerald's contraction, 242
Lyrae β, 262

Machinery, electromagnetic, 109
Machines, 35
Mad Hatter, the, 270
Magnesia, 93
Magnetic, declination, 95; deflexion, 120; dip, 95; flux, 108, 109; force, 93, 94, 96, 97; of a current, 99; induction, 94, 95; moment, 96; permeability, 93, 115, 213; poles, 93, 94
Magnetism, 93
Maize, 178
Malt, 168, 171
Maltase, 171
Maltose, 169, 171
Margarine, 173
Mariner's compass, the, 93
Marsh gas, 10, 11, 154, 159, 160
Mass, 3, 4, 245, 246; action, 146; electric, 233, 245; and inertia, 3, 9, 26; units of, 28; and velocity, 245; and weight, 5, 247
Matter, 1 *et seq.*; conservation of, 6, 65; transmutation of, 229, 234
Maxwell, Clerk, daemon, 77; dielectric medium, 122; electromagnetic waves, 115, 213
Meat, 176
Mechanical equivalent of heat, 63, 64, 105
Melissyl alcohol, 162
Melting point of ice, 153, 127, 129
Mendeléeff's periodic table, 20, 21
Mercury, orbit of, 253, 269
Mercury thermometer, 53, 54, 57
Meta-compounds, 165
Metals, 13; conductivity of, 122; melting points of, 131
Methane, 154, 159, 160
Methyl alcohol, 161
Metre, standard, 23, 24
Metric units, 24, 28
Michelson, velocity of light, 241, 244
Microscope, 205
Milk sugar, 169
Milky Way, 267
Minkowski, space-time, 246
Mirrors, 199, 200
Mist or cloud, 52
Mixed crystals, 133
Molecular, forces, 69, 70, 88; formulae, 154; weights, 45; weight in solution, 139, 140
Molecules, 11, 45, 72, 154 *et seq.*
Moment of inertia, 36
Moments, 36
Momentum, 28
Monatomic gases, 73; solids, 74
Monochlormethane, 154
Monomolecular reaction, 147, 229
Morley, E. W., velocity of light, 241, 244
Moseley, H. G. J., X-ray spectra, 21, 224, 225
Motion, laws of, 27
Motor, electric, 111
Mount Wilson Observatory, 260
Moving coil galvanometer, 101, 105
Muscular contraction, 172
Music, 189

Naphthalene, 165
National Physical Laboratory, 57
Natural paths, 27, 247, 269
Nebulae, 264; irregular, 265; planetary, 265; spiral, 265–267
Nebular theory, 266, 267
Neptune, 253, 257
Neutrality, of blood, 174; of tissues, 174
Neutralization of acids and alkalis, 18, 142, 143
Newland's periodic table, 20
Newton, Sir Isaac, astronomy, 251; definition of mass, 4, 5; dispersion, 206; gravitation, 4, 247, 252; idea of time, 25, 269; laws of motion, 27; pendulums,

5; rings, 196; theory of light, 198; velocity of waves, 187
Nicol's prism, 213
Nitrates, 17, 18, 180
Nitric acid, 17
Nitrogen, 17; disintegration of, 234; in the soil, 180; peroxide, dissociation of, 147; thermometer, 54
Nitroglycerine, 164

Ohm, unit of resistance, 102
Ohm's law, 102
Oil-bearing strata, 159
Oils, vegetable, 173
Olefiant gas, 160
Optical activity, 157; isomers, 158
Organ pipes, 190, 191
Organic, chemistry, 151; compounds, series of, 159; synthesis of, 160, 165, 170, 177
Ortho-compounds, 165
Oscillations, electric, 214, 219, 220; harmonic, 37
Osmotic pressure, 135; of electrolytes, 140
Ostwald, W., dilution law, 146
Oxygen, 7, 8, 12, 45; and hydrogen, 145

Palmitin, 164
Para-compounds, 165
Paraffins, 159
Parallel forces, 34
Parallelogram law, the, 33
Pasteur, L., rotatory power, 157
Pélabon, hydrogen and selenium, 144
Peltier effect, 104
Pendulum, the, 40
Penicillium glaucum, 159
Peptide linkage, 177
Periodic table, 20, 21, 22
Petroleum, 154, 159
Pfeffer, osmotic pressure, 135, 136
Phase, rule, 126; diagram, 128
Phases, equilibrium of, 61, 62, 80, 125, 128
Phenol, 165
Phlogiston, 6, 8, 59
Phosphorus and phosphates, 181
Photography, applied to astronomy, 257
Physical, equilibrium, 124; dimensions, 24, 114; isomers, 156
Phytol, 180
Pigs, food of, 174
Planck, M., Quantum theory, 235–238
Plants, synthetic action of, 179
Platinum, 13; thermometer, 58, 104
Polarization, electrolytic, 98; of light, 211; rotation of the plane of, 157
Pole Star, the, 262
Polonium, 226
Porous plug experiment, 70
Positive rays, 121, 122
Potential, chemical, 126; difference, 101; energy, 35, 124; thermodynamic, 124, 125
Poundal, unit of force, 29
Pressure, of fluids, 45; of vapours, 77, 81, 84, 86
Priestley, J., oxygen, 7
Projectiles, 32, 247
Propane, 155, 160
Propylene, 160
Proteins, 167, 175
Proton, the, 233, 234, 238
Prout's periodic table, 20
Ptolemy, astronomy, 250
Pyrrole, 180
Pythagorean philosophy, 249

Quantum theory, 235–238

Racemic acid, 157
Radiation, atomic, 234; ranges of, in spectra, 224, 258, 263; thermometer, 59
Radicles, 166
Radio-activity, 21, 226 *et seq.*; theory of, 227
Radium, 226; heat evolution of, 229; life of, 229, 232; pedigree of, 230
Rain drops, rate of fall, 2, 51, 52; vapour pressure, 50, 51
Ramsay, Sir William, helium, 208
Rankine, W. J. M., definition of energy, 63

Ratio, of electric units, 115; of specific heats, 68, 69
Rays, α, 227, 231, 233; β, 227, 245; γ, 227; cathode, 119–121; positive, 121, 122; Röntgen or X, 119
Reaction, dimolecular, 148; monomolecular, 147; rates of, 147; reversible, 144, 146
Recombination of ions, 149
Reflexion of light, 198
Refraction of light, 201; index of, 203
Regnault, H. V., specific heats of gases, 65–67
Relativity, 6, 27, 244 *et seq.*; 268–270
Resistance, electric, 102; measurement of, 103, 107; to moving bodies, 2, 31, 51; to ships, 186
Resolution of velocity, acceleration, force, 33
Resonance, 189, 208, 255
Respiration, 7, 175
Reversible reactions, 124, 141, 144, 146
Rigidity, 2, 41
Ring formulae, 165
Ripples, 185
Ritter, stellar evolution, 264
Römer, O., velocity of light, 239
Röntgen, W., rays, 119, 217, 223
Rotation, 36; of the plane of polarization, 157
Rotatory power, 157
Ruhmkorff coil, 112
Russell, H. N., stellar evolution, 264
Rutherford, Sir Ernest, α rays, 232; radioactivity, 227; theory of, 229

Saccharin, 162
Salicylic acid, 166
Salisbury, Marquess of, aether, 238
Salt solutions, 129
Salts, 18; solubility of, 18, 19
Saponification, 149, 164
Sea waves, 185
Secchi, Father, stellar spectra, 262
Selenium and hydrogen, 144
Self-induction, 214
Semi-permeable membranes, 136
Ship resistance, 186
Shipley, Sir Arthur, *Life*, 181
Silk, artificial, 172
Silver, 13, 100
Sirius, 259
Soddy, F., isotopes, 20, 122, 231; theory of radioactivity, 229
Sodium, 14, 207
Soil, 49, 50; nitrogen, 180
Solar, day, 25; system, 256; origin of, 267
Solid solutions, 133
Solids and fluids, 2, 41
Solubility, 134
Solute, 133
Solutions, 129; freezing points of, 138, 139; thermodynamics of, 137; vapour pressure of, 137, 138
Solvent, 133; effect of, 150
Sound-ranging, 189
Sound, shadows, 193; velocity of, 188, 191; waves, 187; interference of, 192, 194
Space, 23, 244 *et seq.*, 268–270
Specific gravity, 42; heat, 60, 236
Specific heats of gases, 65–71; ratio of, 68, 69, 188, 192
Specific inductive capacity, 92
Spectra, 21, 207; absorption, 255; distribution of energy in, 210, 235; enhanced lines, 255; of the stars, 207, 255, 262; types of, 254; X-ray, 224; Zeeman effect, 255
Spectrometer or spectroscope, 207, 251, 254
Stahl, phlogiston, 6
Starch, 169, 170
Stars, the, 259; absolute magnitude of, 260; Cepheid, 262, 268; distance of, 259, 260, 262, 268; double and multiple, 260; giant and dwarf, 260, 263; magnitude of, 260; mass of, 261; number of, 260; parallax of, 259; spectra of, 207, 256, 262; temperature of, 263; universe of, 267; variable, 261, 268
Stationary undulations, 191, 216
Steam calorimeter, 61, 65, 67, 68
Steam engine, 76
Stereo-isomerism, 156

Stokes, Sir G. G., resonance, 208; viscous forces, 52
String recorder, 59
Sugar, 169, 179; inversion of, 149; solutions, 136; synthesis of, 170
Sulphates, 16, 181
Sulphur, 15, 16, 181
Sulphuric acid, 16
Sun, the, 256; absorbing layer of, 258; chromosphere, 258; corona, 259; density of, 257; diameter of, 257; distance of, 256, 257; energy of, 259; mass of, 257; photosphere of, 123; radiating layer of, 258; rotation of, 257; spectrum of, 207, 255, 257; temperature of, 258
Sunlight, action of, 179
Sun spots, 257; spectrum of, 258
Surface, 24; energy or tension, 48–51, 185
Synthesis of organic compounds, 160; of benzene, 165; of peptide-linkages, 177; of sugar, 170; of urea, 177
Synthetic action of plants, 179

Tangent galvanometer, 100
Tartaric acid, 157
Telegraph, electric, 101
Telegraphy, wireless, 217–223
Telephone, 108
Telephony, wireless, 217–223
Telescope, 206
Temperature, 53–59, 72, 74; coefficient of resistance, 104; and radiation, 59, 210, 258, 263
Theory of light, emission, 198; wave, 195, 197, 211, 238
Thermo-dynamic, potentials, 124, 125; scale of temperature, 57–59, 74–76; of solutions, 137
Thermionic valve, 219
Thermionics, 123, 219
Thermo-couple, 58, 104
Thermometer, 53–59
Thermos flask, 86
Thermostats, 130
Thomson, Sir J. J., cathode rays, 119–121; electrons, 232; positive rays, 121
Thomson, William (Lord Kelvin), absolute scale of temperature, 75; and Joule, porous plug experiment, 70
Time, 25, 244 *et seq.*, 269, 270
Tissues, neutrality of, 174
Titration, 19
Torpedo, 239
Torricelli, the, barometer, 46
Tourmaline, 211
Transformer, 112
Transit of Venus, 256
Trichlormethane, 154
Trioxymethylene, 163
Triple point, 127
Tuning fork, 190
Tycho-Brahe, astronomy, 251

Ultra-violet light, 211
Units, electric, 114; ratio of, 115; heat, 59, 60; length, 24; mass, 28; time, 25
Universe, the, 77, 267, 270
Uranium, 226, 228; pedigree of, 230; X, 228
Urea, 177

Vacuum vessel, 86
Valency, 12; unsatisfied, 160, 163
Van der Waals' equation, 87–89
Van't Hoff, osmotic pressure, 135; rotatory power, 157
Vapour pressure, 51, 77–80, 81, 82, 128; of solutions, 137, 138
Vapours, density of, 43, 44
Vegetable oils, 173
Velocity, and acceleration, 25, 26, 30, 31; of light, 115, 116, 213; of sound, 188, 191
Velocity, 25; of the aether, 241; of the ions, 118; of light, 115, 213, 239
Venus, transit of, 256
Viscosity, 51, 52
Vitamines, 178
Volt, unit of E.M.F., 101
Volta, A., electric currents, 21; voltaic cells, 98
Voltmeter, 105
Volume, 24

Waals, Van der, equation, 87–89

Water, composition of, 11, 12; decomposition of, 117; freezing point of, 127; vapour of, 45, 51, 77, 81, 128, 145
Watson, Sir William, electricity, 90
Wave, length, 184, 197, 210; motion, 182; theory of light, 195, 197, 211, 238; front, 216
Waves, 182; canal, 184; electromagnetic, 213, 214, 216; frequency of, 184; refraction of, 201; sea, 185; sound, 187; velocity of, 187
Weight, 5, 7, 247
Wheatstone's bridge, 103, 118
Wilson, C. T. R., cloud tracks, 232; electroscope, 118
Wireless, telegraphy, 123, 217–223; telephony, 123, 217–223
Work, and energy, 34; and power, 29

X-ray spectra, 21, 224, 225
X-rays, 119, 217, 223

Young, T., interference, 195

Zeeman effect, 255
Zein, 178

CAMBRIDGE: PRINTED BY W. LEWIS AT THE UNIVERSITY PRESS